

JEAN-PIERRE CHANGEUX
ALAIN CONNES

MATIÈRE À PENSÉE



EDITIONS
ODILE JACOB

Jean-Pierre Changeux
Alain Connes

MATIÈRE À PENSÉE

ÉDITIONS ODILE JACOB
15, rue Soufflot, 75005 Paris

ISBN 2-7381-0073-2

© ÉDITIONS ODILE JACOB, OCTOBRE 1989

La loi du 11 mars 1957 interdit les copies ou reproductions destinées à une utilisation collective. Toute représentation ou reproduction intégrale ou partielle faite par quelque procédé que ce soit, sans le consentement de l'auteur ou de ses ayants cause, est illicite et constitue une contrefaçon sanctionnée par les articles 425 et suivants du Code pénal.

Avant-propos

Les mathématiciens font en général bon ménage avec les biologistes. Mais ils se parlent peu. Leurs connaissances et leurs motivations sont si éloignées que le dialogue paraît impossible. L'enjeu est cependant considérable. Personne ne niera que l'on fait des mathématiques avec son cerveau. Mais aucune machine construite par l'homme n'est encore parvenue à reproduire les facultés de raison et d'invention de notre machine cérébrale. Y arrivera-t-on un jour? Une authentique intelligence artificielle est-elle réalisable à partir de la matière? Telle est l'interrogation centrale de ce livre.

Avant de répondre à cette question, il est indispensable de définir ce que sont les mathématiques. Quelle est la nature des objets mathématiques? Ceux-ci existent-ils indépendamment du cerveau de l'homme, qui les découvre? Ou, au contraire, sont-ils seulement le produit de l'activité cérébrale, qui les construit? Les développements récents des neurosciences, ou sciences du système nerveux, versent de nouvelles pièces à un dossier instruit déjà au cœur des *Dialogues* de Platon.

Les mathématiques sont les mêmes à Paris, à Moscou et à San Francisco. Mais sont-elles universelles au point de pouvoir nous servir à communiquer avec d'hypothétiques habitants d'autres planètes?... Certes, l'efficacité avec laquelle les mathématiques décrivent le monde qui nous entoure est telle, qu'on la qualifie parfois de déraisonnable. Mais n'est-ce pas là, seulement, l'effet de la fascination qu'exerce, sur son créateur, après coup, l'objet qu'il a créé? Pygmalion mathématicien?

Les réponses à ces interrogations sont, pour la plupart, à rechercher dans l'organisation du cerveau et dans son fonctionnement. Réseau de neurones certes, mais d'une extrême complexité, il doit ses propriétés exceptionnelles à des principes d'architectures et des fonctions élémentaires qu'anatomistes et physiologistes s'efforcent activement d'analyser, et qui, le moment venu, inspireront le constructeur de machines. Mais ces propriétés, le cerveau les doit également à sa nature de système en évolution. Chacun connaît les théories de Darwin sur l'évolution des espèces vivantes. Mais l'idée est peu répandue dans le public que la construction du cerveau durant le développement embryonnaire, puis après la naissance, constitue une évolution au cours de laquelle une sélection s'exerce sur les connexions entre cellules nerveuses. Celle-ci se poursuit par d'autres évolutions, à des niveaux d'organisation plus élevés, qui pourraient rendre compte du déroulement de la pensée, du raisonnement mathématique et pourquoi pas de l'imagination...

Enfin, les connaissances, en mathématiques comme dans les neurosciences, se développent à un point tel que leur impact social devient, chaque jour, plus important. Des problèmes d'éthique se posent. Mais tout d'abord qu'est-ce que l'éthique? La morale peut-elle reposer sur des fondements naturels, qu'il faudrait rechercher dans le fonctionnement du cerveau humain en société? Peut-on fonder une

éthique sur des principes universels semblables à ceux des mathématiques?

La forme de ce livre est celle d'un dialogue. Face à toutes ces interrogations, aucun de nous ne possédait un savoir suffisant de la discipline de l'autre pour se charger, seul, d'avancer des réponses. Mais surtout, le dialogue a permis à chacun d'affiner son point de vue. Nos positions s'accordent sur certains points, divergent sur d'autres, et non des moindres. Mais les questions restent ouvertes, laissant ainsi le troisième partenaire, le lecteur, libre de choisir, et de poursuivre la réflexion, par analogie et par différence avec l'un ou l'autre des protagonistes.

Les *questions d'éthique* qui terminent cet ouvrage nous ont paru appeler une autre forme de présentation que celle du dialogue initial. Le recul de l'écrit s'imposait. Chacun a donc souhaité présenter quelques brèves réflexions écrites, prélude, peut-être, à une tâche future.

Nous remercions Christophe Guias, pour le soin extrême apporté à la transcription des bandes magnétiques, et Jean-Luc Fidel, pour la révision ultime du manuscrit. Enfin, toute notre gratitude va à Odile Jacob pour l'intérêt qu'elle a manifesté, dès l'origine, pour ce dialogue d'idées et pour toutes les facilités qu'elle a très généreusement mises à notre disposition.

Jean-Pierre Changeux et Alain Connes.

Les mathématiques et le cerveau

Platon matérialiste?

La nature habillée sur mesure

Le mathématicien neuronal

Darwin chez les mathématiciens

Les machines à penser

Questions d'éthique

I

Les mathématiques et le cerveau

1. *Présentations*

JEAN-PIERRE CHANGEUX : Avant que nous n'abordions ce qui sera notre première question de fond, la nature des objets mathématiques, je voudrais que nous tentions d'expliquer ce qui nous fait nous tourner ainsi l'un vers l'autre.

Plusieurs points de rencontre entre biologie et mathématiques m'apparaissent. Mon premier contact avec les mathématiques, au lycée et en classe préparatoire aux grandes écoles, fut difficile. La biologie était alors systématiquement dévalorisée par les enseignants, attitude que l'on retrouve d'ailleurs dans les écrits de mathématiciens de très grand renom. On lit par exemple sous la plume de René Thom ¹ que « les progrès de la biologie n'ont pas eu d'effets radicaux quant à l'amélioration de la santé et de la longévité », ou que « les biologistes n'éprouvent pas le besoin

1. R. Thom, *Paraboles et catastrophes*, Flammarion, Paris, 1983, p. 50.

de théorie » ¹. Il y a là une volonté de dévaloriser la biologie, qui s'explique peut-être par la propension des mathématiciens à privilégier la compréhension rapide au détriment d'une réflexion plus lente, plus globalisante, plus imaginative et peut-être plus profonde. Ma réaction a donc été d'abord hostile. Sans doute cachait-elle également le désir de contribuer à l'activité mathématique et de mieux l'assimiler.

C'est seulement dans le cadre de mon travail de recherche en biologie moléculaire, et aujourd'hui en neurobiologie, que j'ai été amené à user très concrètement d'outils mathématiques. Jacques Monod a été un maître exceptionnel à cet égard. Avec lui, j'ai pu développer plusieurs modèles en biologie moléculaire, en particulier en ce qui concerne les protéines allostériques ², molécules spécialisées dans la régulation. Dans ce cas précis, les mathématiques nous ont permis de donner forme à nos idées et de développer des prédictions quantitatives. Aujourd'hui, dans mon travail de neurobiologiste, l'outil mathématique me paraît toujours plus nécessaire à la construction de modèles rationnels des fonctions cérébrales. Un domaine de recherche très nouveau se développe d'ailleurs aux confins des neurosciences, de la psychologie et des mathématiques : les sciences cognitives. Leurs progrès, d'ores et déjà, paraissent dépendre d'une étroite collaboration entre théoriciens et expérimentateurs.

Plus généralement, ce qui me pousse à m'intéresser étroitement aux mathématiques, c'est le besoin de comprendre comment le cerveau crée et utilise les objets mathématiques, quelles sont les relations entre les mathématiques et le cerveau. A elle seule, cette question justifie notre rencontre.

1. *Ibid.*, p. 46.

2. J. Monod, J. Wyman et J.P. Changeux, *On the nature of allosteric transition : a plausible model*, J. Mol. Biol. n° 12, 1965, p. 88-118.

Mais les mathématiques jouent également un rôle central dans la vie sociale. La culture occidentale se caractérise par une sorte de mythe des mathématiques : la croyance, peut-être issue de Pythagore, en une vertu explicative et presque transcendante des mathématiques. Pour beaucoup, décrire en termes mathématiques une structure syntaxique ou des relations de parenté paraît une « explication » suffisante. Plus pratiquement, l'ordinateur et ses applications confèrent aux mathématiques un pouvoir unique, toujours croissant. Le récent krach de Wall Street n'est-il pas en partie dû au « comportement programmé » d'ordinateurs qui agissaient pour le meilleur « bénéfice » de leurs clients ? L'ordinateur semble remplacer le cerveau... sans pourtant en avoir les performances ! Ce problème, marginal par rapport à nos activités scientifiques, doit nous amener à réfléchir sur les rapports des mathématiques et de l'éthique, et à nous demander notamment s'il est possible de fonder une éthique universelle des sociétés humaines qui repose sur la rigueur des mathématiques. Cette démarche est-elle complémentaire de la recherche des bases neurales de l'éthique, ou s'en distingue-t-elle radicalement ? Telles sont mes motivations de biologiste. Quelles sont les tiennes ?

ALAIN CONNES : Je répondrai en t'avouant mon enthousiasme face à la question des relations entre les mathématiques et le cerveau et de la nature des objets mathématiques.

Lorsque tu parlais de l'opposition institutionnelle des mathématiques et de la biologie, tu as cité René Thom. C'est certainement un penseur original. Mais il serait dangereux de le considérer comme le porte-parole de l'opinion des mathématiciens. Parlons plutôt d'Israël Gelfand. Son influence sur les mathématiques est considérable. Or, il consacre une grande partie de son activité scientifique à la biologie. Plus de la moitié de ses articles sont consacrés à

cette discipline, et il dirige deux séminaires, l'un de mathématiques, l'autre de biologie.

Quant à moi, c'est la lecture de ton livre, *L'Homme neuronal*, qui m'a fait comprendre que le fonctionnement du cerveau est désormais connu avec une certaine précision. J'ai surtout été frappé par l'existence de cartes perceptuelles, beaucoup plus nombreuses chez l'homme que chez d'autres animaux. Elles relient la rétine à des domaines du cerveau aux fonctions d'interprétation différentes. J'ai également été frappé par les expériences de Shepard¹. Lorsqu'on demande à un sujet de reconnaître si deux objets procèdent l'un de l'autre par rotation dans l'espace à trois dimensions, l'expérience montre que le temps de réponse est proportionnel à l'angle de rotation. Le fonctionnement cérébral obéit ainsi aux lois de la physique. Mais il me semble important de dépasser le domaine particulier de la biologie pour étudier le cerveau. Pour ce faire, les mathématiques fournissent un terrain beaucoup plus propice que d'autres. Parce qu'elles sont absolues, universelles, et donc indépendantes de toute influence culturelle.

JPC : Tu t'engages vers une prise de position...

AC : Il me semble que les notions que chaque langue exprime dépendent de données mal définies, parce qu'influencées par la culture. Au contraire, les objets mathématiques – et c'est ce que je voudrais démontrer – ont une pureté beaucoup plus grande. Ils sont dégagés de cette gangue culturelle, et doivent donc permettre de mieux tester notre compréhension du fonctionnement du cerveau.

Mais mon approche, évidemment, est intéressée. J'aimerais en savoir plus sur la biologie, pour en tirer des enseignements. Ton livre m'a amené à réfléchir sur la manière dont le cerveau assimile une nouvelle théorie, ou

1. R. Shepard et J. Metzler, *Mental rotation of three-dimensional objects*, Science n° 171, 1971, p. 701-703.

bien se familiarise avec une nouvelle activité, comme le jeu d'échecs ou le piano. J'ai dû réviser certaines idées toutes faites sur l'apprentissage, ou corriger certaines erreurs. Par exemple, quand un mathématicien travaille sur un domaine ni trop difficile ni trop étendu, il se peut qu'il parvienne à maîtriser une technique précise. Les mathématiques étant très abstraites, il peut croire cette maîtrise éternelle, et avoir le sentiment qu'il n'est plus nécessaire de travailler pour pouvoir la retrouver à tout moment. Comme ton livre m'a permis de le comprendre, ce savoir-faire est probablement localisé dans une zone précise du cerveau : si le système de neurones correspondant n'est pas excité de temps en temps par l'utilisation de cette technique, il dépérit.

JPC : Donc, il existe une matérialité de la trace constituée par l'expérience mathématique passée.

AC : Exactement. Il est nécessaire, de temps en temps, d'ouvrir un tiroir fermé depuis des années. Sinon l'inutilité apparente de son contenu entraîne sa destruction progressive.

2. La hiérarchie des sciences mise en cause

JPC : Je souhaiterais que notre discussion aborde trois thèmes : tout d'abord la relation des mathématiques avec les autres sciences, puis la question du réalisme et du constructivisme, et enfin celle du rapport des nombres et de l'expérience.

Quant au statut des mathématiques par rapport aux autres sciences, deux attitudes s'opposent : celle de Descartes et Leibniz, et celle de Diderot. Pour les premiers, les mathématiques éclairent le monde de leur vérité et permettent d'unifier l'ensemble des sciences. Quel que soit l'objet étudié,

il finit toujours par se ramener à du mathématique! Il existe donc une hiérarchie des sciences qui constitue, encore aujourd'hui, le fondement de notre système éducatif. Diderot, bien que proche de mathématiciens aussi distingués que d'Alembert, rejette ce présupposé. Selon lui, les mathématiques n'ajoutent rien à l'expérience : elles ne font qu'« interposer un voile entre la Nature et le peuple »! Francis Bacon, en 1623, écrivait déjà : « Car je ne sais comment il se fait que la logique et les mathématiques, qui ne devraient être que les servantes de la physique, se targuant parfois de leurs certitudes, veulent absolument lui faire la loi ¹. »

AC : Il est courant et relativement justifié de considérer les mathématiques comme un langage nécessaire à la formalisation de presque toutes les autres sciences. Que cette formalisation soit quantitative ou qualitative, elle se fera toujours à travers les mathématiques.

JPC : C'est un peu ce que pensent Descartes et Leibniz.

AC : Oui, mais ils ajoutent que tout finit toujours par se ramener aux mathématiques. Une histoire bien connue des physiciens suggère justement l'inverse. Un physicien qui assistait à une conférence depuis une semaine avait accumulé beaucoup de linge sale. Il se mit à la recherche d'une laverie. Au bout d'un moment, se promenant dans la grande rue de la ville, il aperçut une boutique dont l'enseigne annonçait « Epicerie-Boulangerie-Laverie ». Il entra avec son paquet de linge sale et demanda quand il pourrait être prêt. Le mathématicien qui tenait la boutique lui répondit : « Nous sommes désolés, mais nous ne lavons pas le linge ». « Mais, s'étonna le physicien, j'ai pourtant bien lu " Laverie " sur votre devanture! » Et le mathématicien de répondre : « Nous ne lavons rien... nous vendons seulement des enseignes! » Le physicien repartit donc et lava son linge lui-même. Les mots, comme le montre cette histoire, ne suffisent pas! Les

1. F. Bacon, *De dignitate et augmentis*, III, 6, trad. Buchon, p. 103.

physiciens utilisent les mathématiques comme un langage, mais le contenu effectif de leur science ne se laisse pas réduire aux seules mathématiques.

JPC : Les mathématiques sont un langage plus rigoureux, ni plus ni moins.

AC : Mais un article de physique ne se réduit pas à sa seule expression mathématique. Le physicien utilise souvent des hypothèses qu'il ne précise pas et qui ont pour origine ce qu'on appelle « l'intuition physique ». Elles lui permettent, en particulier, de négliger certaines quantités ou de faire des approximations que le mathématicien aurait bien du mal à deviner. Par exemple, il a fallu une vingtaine d'années, entre 1930 et 1950, pour que les physiciens arrivent à élaborer la méthode de *renormalisation* en théorie des champs. Elle consiste à effectuer un calcul perturbatif dont tous les termes, à partir du deuxième ordre, donnent des intégrales divergentes. Les physiciens¹, motivés par l'extraordinaire précision des résultats expérimentaux de la spectroscopie de la fin des années 40 (structure fine des raies du spectre d'émission des atomes)², ont désespérément cherché à tirer un résultat fini de ces intégrales divergentes. Pour cela, ils ont restreint le domaine d'intégration aux énergies de l'ordre de MC^2 , où M est la masse de l'électron et C la vitesse de la lumière. Moyennant des soustractions non justifiées, ils ont obtenu un résultat fini qui se rapproche beaucoup du résultat expérimental. Cette technique a été progressivement améliorée, par Tomonaga, Schwinger, Feynman et Dyson, jusqu'à un accord avec les résultats expérimentaux qui correspond à l'épaisseur d'un cheveu sur la distance Paris-New York. Quel fut le rôle de l'intuition physique dans leur raisonnement? Le mécanisme de renor-

1. W.A. Bethe, *The Electromagnetic Shift of Energy Levels*, Physical Review vol. 72, p. 339, 1947.

2. Lamb, Retherford, *Fine Structure of the Hydrogen Atom by a Microwave Method*, Phys. Rev. 72, p. 241, 1947.

malisation consiste, au cours des calculs, à changer la masse de l'électron et à la remplacer par une quantité qui dépend de l'ordre de grandeur des énergies considérées, mais diverge quand l'ordre de grandeur tend vers l'infini. Pour prendre une comparaison très simple, si un ballon gonflé d'hélium quitte le sol à un instant $T = 0$, le calcul de son accélération par la poussée d'Archimède ne donnera pas le résultat que l'on observe expérimentalement. En effet, la présence d'un champ, l'air ambiant, équivaut à remplacer dans les calculs la masse réelle du ballon par une masse effective beaucoup plus grande. S'appuyant sur cette comparaison, on peut comprendre que l'électron, placé dans le champ électromagnétique, possède une masse effective bien différente de sa masse « réelle », c'est-à-dire de celle qui rentre dans l'équation mathématique. Grâce à cette intuition, les physiciens ont pu développer une méthode, celle de la renormalisation, qui, bien sûr, se formule en langage mathématique, mais que les mathématiciens, confrontés au même problème, auraient été bien en peine de découvrir. Cette intuition physique les autorise d'ailleurs à prendre des libertés avec la rigueur mathématique. Par exemple, l'intégrale de Feynman ne correspond pour le moment à aucun objet mathématique précis. C'est pourtant le pain quotidien des physiciens théoriciens.

Pourtant on aurait tort de croire que les mathématiques ne jouent vis-à-vis de la physique que le rôle de langage exprimant des résultats. Lorsqu'on modélise une théorie qui en est à un stade assez primitif, les mathématiques ont bel et bien cette fonction. Mais à un stade ultérieur d'élaboration, comme dans le cas de la mécanique quantique, le caractère génératif des mathématiques finit par jouer un rôle crucial. Comment ne pas être troublé par la possibilité de retrouver le tableau périodique des éléments de Mendelév à partir de l'équation de Schrödinger et du principe d'exclusion de Pauli? C'est pourquoi le mathé-

maticien peut croire réduire la physique à un certain nombre d'équations. Bien souvent, cependant, c'est l'intuition du physicien qui lui permet de comprendre ces équations.

JPC : Tu veux dire que c'est le contexte expérimental, en physique, qui permet de créer les objets mathématiques. Une équation ne tombe pas du ciel un beau jour. Elle s'insère dans l'histoire des relations du physicien avec son objet. Progressivement, celui-ci forge un outil mathématique adapté au problème qu'il se pose.

AC : Ce n'est pas tout. Un mathématicien peut arriver à manipuler des objets qui ont une signification physique. Mais s'il n'est pas pleinement conscient de la manière dont, historiquement, ces objets ont été introduits, il risque de commettre très facilement des erreurs que le physicien, lui, ne commet pas. Dire que les mathématiques forment un langage qui contient *exactement* ce que les physiciens ont trouvé constitue une forme d'autoritarisme exagéré. Les physiciens renâclent à exprimer leur point de vue de manière suffisamment précise mathématiquement, de peur de l'appauvrir. A l'inverse, certains développements récents ¹ dans l'interprétation de la mécanique quantique montrent que l'effort de formalisation mathématique peut permettre d'éviter des paradoxes souvent dus à une inadéquation du langage utilisé par les physiciens ou à un manque de réflexion sur la logique elle-même.

JPC : Donc, le langage mathématique est un authentique langage. Mais est-ce le seul ?

AC : C'est le seul langage universel. Indéniablement. Pour le comprendre, imaginons comment on ferait pour communiquer avec une autre intelligence, une autre planète ou un autre système solaire... Il est bien évident que ces

1. R. Omnes, *From Hilbert Space to Common Sense : A synthesis of Recent Progress in the Interpretation of Quantum Mechanics*, Preprint, Lab. Phys. théorique et Haute énergie, Orsay, 89/7.

« gens-là » ne parleraient aucun des langages que nous pratiquons, et qu'ils ne vivraient pas dans une atmosphère formée d'un mélange d'oxygène et d'azote, véhicule de la parole.

JPC : Mais pour que nous puissions communiquer avec eux, il faudrait qu'ils aient la même mathématique que nous ?

AC : J'en suis persuadé... Je pense même que les mathématiques seraient le meilleur moyen de communiquer avec eux. Nous leur communiquerions la liste des nombres entiers, disons de 1 à 100. Nous enverrions le signal suivant : un « top », un long silence, suivis de deux « top » et d'un long silence, puis trois « top », suivis d'un long silence, et ainsi de suite. Une fois cette liste donnée, nous leur communiquerions la loi de l'addition. La seule variable sur laquelle on peut moduler est le nombre de « top » et l'intervalle de temps qui les sépare les uns des autres. Par exemple, pour communiquer $3 + 2 = 5$, le message serait : trois « top » consécutifs, un silence, deux « top » consécutifs, un silence double, et cinq « top ». Il s'agirait bien sûr de faire en sorte que le message ne soit pas ambigu. Il serait ainsi possible de leur communiquer la table d'addition et la table de multiplication dans des limites raisonnables. La difficulté principale consiste à s'assurer qu'ils ont compris. Pour cela, on pourrait par exemple leur envoyer une addition incomplète. Il est probable qu'il faille alors attendre des millénaires avant d'obtenir la réponse ! Il n'empêche qu'une réponse positive serait une preuve indéniable de l'existence d'une autre intelligence au-delà de notre système solaire. Une preuve plus solide que les signaux périodiques issus de l'espace interstellaire, comme ceux qui étonnèrent les astronomes lorsqu'ils découvrirent les premiers pulsars. A un niveau plus élevé, nous pourrions alors leur communiquer la suite des nombres premiers, disons de 1 à 1000, et leur demander le suivant.

JPC : On risque d'attendre longtemps la réponse avant de trancher. Et quand bien même cette communication serait établie, que prouverait-elle ? Tu affirmes que « ces gens-là ne parleraient aucun des langages que nous pratiquons », mais qu'« ils utiliseraient la même mathématique ». Je crains de ne pas être d'accord. Car il est probable que plusieurs processus cérébraux fondamentaux sont, chez l'homme, communs à l'usage de tout langage, y compris le langage mathématique. Si ces extra-terrestres utilisaient des « mathématiques humaines », ils posséderaient un système nerveux, un cerveau, très proche de celui de l'homme !

3. *Invention ou découverte ?*

JPC : Venons-en à la nature des objets mathématiques. Deux positions diamétralement opposées ont été défendues, le « réalisme » et le « constructivisme ». Pour le « réaliste », qui s'inspire directement de Platon, le monde est peuplé d'Idées qui ont une réalité distincte de la réalité sensible (*voir figure 1*). Nombreux sont les mathématiciens contemporains qui se jugent « réalistes ». Dieudonné, par exemple, écrit dans son livre : « Il est assez difficile de décrire les idées de ces mathématiciens qui varient d'ailleurs de l'un à l'autre. Ils admettent que les objets mathématiques possèdent une “réalité” distincte de la réalité sensible (peut-être semblable à celle que Platon accordait à ses “Idées” ?). » Un mathématicien aussi distingué que Cantor a écrit que « La plus haute perfection de Dieu est la possibilité de créer un ensemble infini et son immense bonté le conduit à le créer. » Nous voilà en pleine *mathesis divina*, en pleine métaphysique ! Ce qui surprend chez des scientifiques sérieux. Descartes déjà se référerait à la métaphysique à propos de la



FIGURE 1

Gravure du *xvii^e* siècle illustrant un passage célèbre de *La République* de Platon sur l'allégorie de la Caverne. Socrate et Glaucon s'interrogent sur la « réalité » des ombres projetées sur les murs de la caverne, comparée à celle



des objets qui les causent. Pour Platon, l'apparence n'est que l'ombre de la réalité, et les Idées ont une existence indépendante du reste de l'univers. (Bibliothèque nationale, Estampes; cliché Jean-Loup Chasuret).

géométrie : « ...lorsque j'imagine un triangle, écrivait-il, encore qu'il n'y ait peut-être en aucun lieu du monde hors de ma pensée une telle figure, et qu'il n'y en ait jamais eu, il ne laisse pas néanmoins d'y avoir une certaine nature ou forme, ou essence déterminée de cette figure, laquelle est immuable et éternelle, que je n'ai point inventée et qui ne dépend en aucune façon de mon esprit ¹ ». Pour les « constructivistes », les objets mathématiques sont des êtres de raison qui n'existent que dans la pensée du mathématicien. Et non dans un monde platonicien indépendant de la matière. Ils n'existent que dans les neurones et les synapses des mathématiciens qui les produisent comme de ceux qui les comprennent et les emploient. On retrouve ce point de vue, évidemment poussé à l'extrême, chez les philosophes empiristes, comme Locke ou Hume. Celui-ci écrit, par exemple, que « toutes nos idées sont copies de nos impressions ». Pour lui, les objets de la géométrie proviennent exclusivement de l'expérience. Comment te situes-tu par rapport à ces deux points de vue opposés ?

AC : Je pense être assez proche du point de vue réaliste. Pour moi, la suite des nombres premiers, par exemple, a une réalité plus stable que la réalité matérielle qui nous entoure. On peut comparer le mathématicien au travail à un explorateur à la découverte du monde. La pratique découvre des faits bruts. On s'aperçoit par exemple, en faisant des calculs simples, que la suite des nombres premiers ne semble pas avoir de fin. Le travail du mathématicien consiste alors à démontrer qu'il existe une infinité de nombres premiers. C'est un vieux résultat qu'on doit à Euclide. Cette démonstration prouve que si quelqu'un affirme un jour avoir trouvé le plus grand nombre premier, il sera

1. R. Descartes, *Méditations métaphysiques*, V, Gallimard, coll. Pléiade, 1970, p. 311.

facile de lui montrer qu'il a tort. On se heurte donc à une réalité aussi incontestable que la réalité physique.

Dans sa quête de la réalité mathématique, le mathématicien crée des « outils de pensée ». Il ne faut pas les confondre avec la réalité mathématique elle-même. Par exemple, le système décimal est un outil de pensée familier, mais on aurait tort d'attribuer une signification aux chiffres décimaux qui apparaissent dans un nombre. Nous allons bientôt fêter l'an 2000. L'importance de ce nombre est pourtant seulement un phénomène culturel. En mathématique, le nombre 2000 est dépourvu d'intérêt. Mais parmi les méthodes dont dispose le mathématicien pour explorer la réalité mathématique, je pense surtout à l'axiomatique. Elle sert à poser les problèmes de classification d'objets mathématiques définis par des conditions très simples. On sait ainsi, par exemple, déterminer exactement la liste de tous les corps finis. Un corps fini est un ensemble fini muni d'une loi d'addition et de multiplication pour laquelle tout nombre non nul a un inverse. Les règles vérifiées par l'addition et la multiplication sont les mêmes que les règles familières pour l'addition et la multiplication des nombres entiers. On démontre qu'il existe, pour chaque nombre premier « p » et chaque entier « n », un corps fini, et un seul, ayant p^n éléments. Disposer d'un théorème de ce genre nous assure qu'une région des mathématiques a été explorée dans ses moindres recoins, tout au moins quant à la liste de ses objets possibles, et ce, sans support matériel.

JPC : Il me semble au contraire que ces objets mathématiques existent matériellement dans ton cerveau. Tu les examines intérieurement par un processus conscient, au sens physiologique du terme. Si tu peux en étudier les propriétés, c'est parce que ces objets ont une réalité matérielle. Tu as mentionné le cas des rotations mentales ¹ et des objets

1. R. Shepard et J. Metzler, 1971, réf. p. 18.

que notre cerveau traite de manière physique. Notre cerveau est un objet physique complexe. Comme tel, il construit des « représentations » identifiables à des états physiques. Les objets mathématiques seraient, dans la tête du mathématicien, des objets matériels, des « objets mentaux »¹ aux propriétés analysables par un processus réflexif. Celui-ci peut bien faire appel à d'autres objets mathématiques plus banals, que tu appelles « outils ». Mais je ne les considère pas comme de nature radicalement différente, bien que de niveaux de complexité ou d'abstraction différents. Enfin, le travail mathématique requiert des facultés cérébrales de raisonnement, de logique, qui me semblent directement liées à l'organisation de notre cerveau, et qui existaient déjà, au moins en partie, dans celui-ci lorsque l'*Homo erectus* développait ses stratégies de taille d'outil de pierre (voir figure 2). Ces « objets mathématiques » s'identifient à des états physiques de notre cerveau de telle sorte qu'on devrait *en principe* pouvoir les observer de manière extérieure grâce à des méthodes d'imagerie cérébrale. Leur résolution est encore trop insuffisante pour qu'on puisse effectivement y parvenir, mais l'idée se défend.

AC : Si l'on accepte l'existence d'une réalité mathématique indépendante de l'homme, il faut nettement distinguer cette réalité et la manière dont elle est appréhendée. Il est clair que, pour la percevoir, notre cerveau utilise une imagerie cérébrale proche de la physique, du moins pour la géométrie ordinaire fondée sur les nombres réels et l'espace euclidien. Cependant, la méthode axiomatique, pour ne citer qu'elle, permet au mathématicien de s'aventurer bien au-delà de cette contrée familière. Comment l'imagerie mentale fonctionne-t-elle dans ces régions-là ? Prenons un exemple. On arrive à classifier complètement les corps localement compacts. On sait déterminer exactement les corps,

1. J.-P. Changeux, *L'homme neuronal*, Fayard, Paris, 1983.

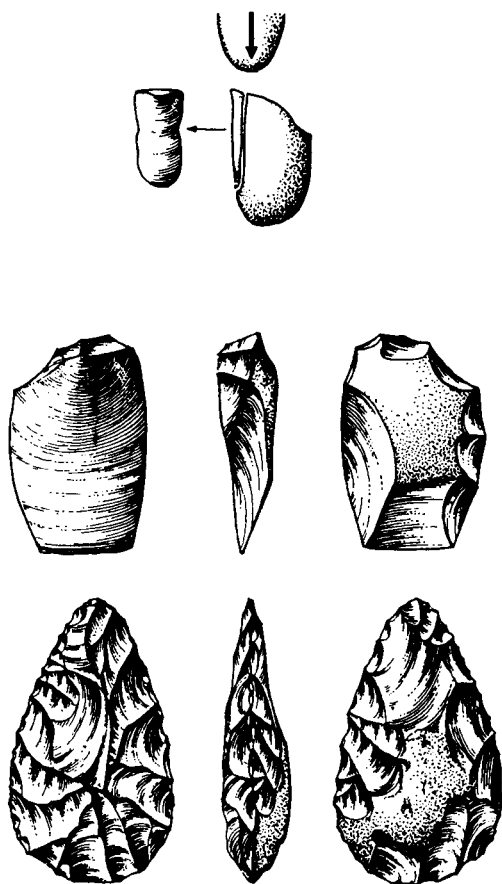


FIGURE 2

Les ancêtres de l'Homo sapiens développent des stratégies de taille d'outils de pierre qui exigent à la fois la maîtrise du geste et un taux élevé de prévision dans le déroulement des opérations manuelles. Les facultés de représentation et de raisonnement logique étaient déjà très développées chez l'Homo erectus qui a confectionné ces outils et qui, il y a 400 000 ans environ, domestiqua le feu... Une légère asymétrie entre les empreintes des hémisphères gauche et droit sur les os du crâne suggère que l'Homo erectus avait déjà l'usage de la parole. (D'après Leroi-Gourhan, A., Le geste et la parole, Albin Michel, Paris, 1964)

c'est-à-dire les objets mathématiques dans lesquels on a une loi d'addition et de multiplication, où tout nombre non nul a un inverse, et qui sont localement compacts. On connaît la droite réelle, qui sert de support à la physique. Mais il y a aussi ces corps très étranges qu'on appelle péadiques (*voir figure 3*). Pour le moment, jamais ils n'ont servi à résoudre un seul problème de physique. Mais ils existent, et ils sont paramétrés par un nombre premier, de sorte qu'à chaque nombre premier correspond un corps péadique. On connaît également de petites variations, qu'on appelle des extensions algébriques : le corps des nombres complexes, puis les extensions algébriques des corps péadiques, et enfin des corps de séries formelles sur les corps finis. Parmi tous ces corps, un seul, ou plutôt deux, les réels et les complexes, ont été utilisés en physique. On peut faire des calculs avec des nombres comme les nombres péadiques. Mais tout se passe comme si, au lieu de calculer de gauche à droite, on calculait de droite à gauche. La notion de taille ou de grandeur d'un nombre ne correspond plus à la notion ordinaire. Ces calculs peuvent être effectués par un ordinateur aussi bien que par le cerveau humain. Mais il est difficile de trouver un modèle physique simple qui serve d'imagerie mentale à ces calculs. Je pense que les possibilités d'adaptation du cerveau lui permettent justement de développer une intuition, qui ne provient pas de la réalité physique, mais qui est adaptée au problème mathématique posé.

JPC : Il me semble que tu ne distingues pas suffisamment les objets mathématiques eux-mêmes de leurs propriétés. Ces objets sont des « constructions nouvelles » que le mathématicien conçoit avant d'en avoir examiné toutes les propriétés. Au départ, ce sont des « conjectures », des « postulats », qui peuvent être ou non démontrés. C'est dans la conjecture, dans l'édifice initial postulé, qu'on touche à la nature des objets mathématiques. John Stuart Mill proposait déjà de dire que « le fait énoncé dans la *définition*

Nombre réel en écriture dyadique	10,010110100010...
Addition de deux nombres réels	$ \begin{array}{r} 10,0101101000100... \\ + 1,1001001100110... \\ \hline = 11,1110110101010 \end{array} $

...0010001011010,01	Nombre p-adique, p=2
$ \begin{array}{r} ...0010001011010,01 \\ ...0110011001001,1 + \\ \hline ...1000100100011,11 \end{array} $	Addition de deux nombres p-adiques, p=2

FIGURE 3

Cette figure montre sur un exemple la comparaison entre l'addition de deux nombres réels en écriture dyadique et l'addition de deux nombres 2-adiques. La connexité des nombres réels résulte des identifications de la forme suivante : $0,00111111... = 0,0100000...$

d'un nombre est un fait physique » ¹. Rien d'étonnant à ce que les nombres entiers possèdent telle ou telle propriété. Celle-ci est contenue dans la définition que le mathématicien propose, dans l'intuition de départ. Mais il faut du temps pour reconnaître ces propriétés. L'axiomatique, la logique, et toutes les fonctions cérébrales qui s'y rapportent, jouent alors un rôle crucial dans ce travail d'*analyse*, de *déduction* : elles servent d'*appareil logique*. Un des traits les plus frappants de la machine cérébrale humaine est donc tout autant de créer des objets mentaux nouveaux que d'en analyser les propriétés qui paraissent souvent, mais *a posteriori*, d'une extrême simplicité.

AC : En classe élémentaire, on apprend aux enfants à faire des additions et des divisions à partir de nombres réels. Il serait beaucoup plus délicat de leur apprendre à manipuler les nombres péadiques. Pourquoi ? Parce qu'ils devraient avoir franchi un cap très important dans la pratique des mathématiques : celui du contact avec le réel. Au-delà, on perd le sens immédiat des grandeurs, on doit se livrer au seul calcul. La réalité à laquelle on se heurte n'est plus la réalité tangible d'un triangle isocèle ou non. Elle est bien plus forte. Si l'on fait un calcul de deux manières différentes, et si l'on ne trouve pas le même résultat, on éprouve une réelle frustration. Pour moi, la réalité mathématique est de cet ordre. Il existe une cohérence, inexpliquée justement, indépendante de notre système de raisonnement, qui garantit que si l'on travaille correctement, on trouvera l'erreur. On découvre alors une cohérence qui dépasse vraiment celle que produit l'intuition sensible, l'intuition directe des phénomènes.

JPC : Que cette cohérence soit encore inexpliquée ne prouve pas qu'elle est *inexplicable*. Encore moins qu'elle est

1. J. S. Mill, *Système de logique*, III, 24.5, 1851.

indépendante de notre système de raisonnement, comme tu l'affirmes.

4. *Les mathématiques ont une histoire*

JPC : Je reste donc dubitatif vis-à-vis de l'opinion suivant laquelle les objets mathématiques existent quelque part « dans l'Univers » indépendamment de tout support matériel et cérébral. Il me paraît utile de prendre certaines distances vis-à-vis du travail du mathématicien et, en particulier, vis-à-vis des objets qu'il construit. Il faudrait situer l'objet mathématique dans le contexte historique où il est apparu. On enseigne les mathématiques comme un ensemble cohérent de propositions, de théorèmes, d'axiomes. On oublie qu'ils sont apparus progressivement au cours de l'histoire des mathématiques et des sociétés humaines, bref qu'il s'agit d'objets culturels sujets à évolution. Replacer les objets mathématiques dans une perspective historique permet au contraire de les « laïciser », de les rendre plus contingents qu'ils ne paraissent. Les théories succèdent aux théories, et certaines, sans infirmer les précédentes, apportent un nouvel éclairage. C'est le cas par exemple des géométries non euclidiennes. Les axiomes de la géométrie euclidienne forment un tout cohérent : on est bien en présence de cette fameuse cohérence qui t'étonne tant et qui rend l'ensemble, mais *en apparence* seulement, indépendant de tout support matériel, pour reprendre tes termes. Néanmoins, au XIX^e siècle, les géométries non euclidiennes sont venues tout perturber.

AC : Mais elles n'ont en rien perturbé la cohérence de la géométrie euclidienne ! On peut au contraire utiliser cet exemple pour montrer la puissance et la fécondité de l'outil axiomatique. Au départ, la géométrie euclidienne a été

comprise par l'intermédiaire de l'expérience physique. Euclide a essayé de poser un certain nombre d'axiomes qui permettent d'effectuer ce qu'on appelle des démonstrations. L'un d'eux semblait tout à fait superflu : l'axiome de l'unique parallèle à une droite donnée passant par un point donné. Il paraissait possible de démontrer qu'il n'était pas nécessaire d'ajouter celui-là et qu'il découlait des autres. C'est justement à force d'essayer d'en démontrer la nécessité, qu'on a découvert les géométries non euclidiennes. Pendant une bonne partie du XIX^e siècle, celles-ci ont été considérées comme ésotériques par les mathématiciens. Gauss a même hésité à publier ses résultats, de peur de susciter l'incrédulité. Jusqu'au jour où Poincaré s'est aperçu que la géométrie plane de courbure -1 était un outil extraordinaire, même pour résoudre des problèmes de théorie des nombres qu'il avait considérés indépendamment. Il en a tiré sa théorie des fonctions fuchsiennes. Comment est-on donc arrivé aux géométries non euclidiennes? Non pas parce que nous avons constaté que l'espace dans lequel nous vivons n'est pas conforme à la géométrie euclidienne. Mais simplement à partir d'un problème axiomatique et de l'effort pour caractériser la géométrie par un petit nombre de propriétés.

JPC : Cela ne prouve toujours pas l'immatérialité des objets mathématiques! Pour moi, la méthode axiomatique est l'expression, élaborée de facultés cérébrales, de facultés cognitives, liées à l'usage du langage chez l'homme. Or ce qui caractérise le langage, c'est précisément son caractère *génératif*.

AC : Intervient ici une caractéristique, propre aux mathématiques, très malaisée à expliquer. On peut souvent, moyennant des efforts considérables, arriver à une liste complète d'objets mathématiques définis par des conditions très simples. On croit intuitivement que la liste est complète, et on cherche en général à démontrer qu'elle est exhaustive. Or, souvent, on trouve d'autres objets, précisément en cher-

chant à démontrer qu'on a épuisé la liste. Prenons l'exemple de la théorie des groupes finis. La notion de groupe fini est élémentaire et presque de même niveau que celle de nombres entiers. Un groupe fini est le groupe des symétries d'un objet fini. Les mathématiciens ont cherché à classifier ce qu'on appelle les groupes finis simples, c'est-à-dire les groupes finis qui, un peu comme les nombres premiers, ne peuvent se décomposer en groupes plus petits. C'est un problème extrêmement difficile. Galois avait montré que pour $n \geq 5$ le groupe des permutations paires d'un ensemble à n éléments est simple. Et le Français Claude Chevalley avait construit des séries de groupes finis simples, qui ressemblaient à ce qu'on appelle les séries de groupes de Lie. On pouvait alors penser qu'outre les groupes découverts par Mathieu au siècle dernier, il n'en existait pas d'autres. Lorsqu'on a cherché à le démontrer, on a trouvé une vingtaine de groupes qui n'étaient pas compris dans la liste de Chevalley : les groupes sporadiques. Il y a une quinzaine d'années, on a découvert le dernier groupe fini simple, celui qu'on appelle « le monstre ». C'est un groupe fini, découvert par raisonnement purement mathématique, dont le nombre d'éléments est considérable :

808017424794512875886459904961710757005754368000000000

Aujourd'hui, les spécialistes ont réussi, au prix d'efforts héroïques, à démontrer que la liste des 26 groupes finis simples sporadiques est bien complète (*voir figure 4*).

JPC : Je ne vois pas en quoi épuiser toutes les possibilités démontre que l'objet en question est une « idéalité » qui préexiste à l'homme. Prenons l'exemple d'un objet régulier, comme un cube ou une pyramide de sel gemme. Il est évident que ses propriétés vont être rapidement épuisées. Cela ne prouve pas, bien que Descartes l'affirme, que ses propriétés sont celles d'une forme « immuable et éternelle »

LES GROUPES FINIS

Définition

Un groupe fini G est la donnée d'un ensemble fini G et d'une loi de composition, i.e. une application de $G \times G$ dans G notée $(g_1, g_2) \rightarrow g_1 g_2$ telle que : 1) $g_1 (g_2 g_3) = (g_1 g_2) g_3$ pour tous $g_i \in G$; 2) il existe $e \in G$, tel que $eg = ge = g$; $\forall g \in G$; 3) pour tout $g \in G$, il existe $g^{-1} \in G$ tel que $gg^{-1} = g^{-1}g = e$. Un homomorphisme d'un groupe G_1 dans un groupe G_2 est une application l de l'ensemble G_1 dans l'ensemble G_2 telle que $l(g_1 g_2) = l(g_1) l(g_2)$, $\forall g_1, g_2 \in G_1$.

Un groupe fini G est simple si et seulement si tout homomorphisme de G vers G' est soit constant soit injectif.

Théorème de classification des groupes finis simples

Les groupes finis simples sont les suivants :

- les groupes cycliques d'ordre premier
- les groupes alternés de degré 5 au moins
- les groupes de Chevalley et le groupe de Tits
- les 26 groupes sporadiques

Les groupes sporadiques

GRUPE	ORDRE	INVESTIGATEUR
M_{11}	$2^4.3^2.5.11$	Mathieu
M_{12}	$2^6.3^3.5.11$	Mathieu
M_{22}	$2^7.3^3.5.7.11$	Mathieu
M_{23}	$2^7.3^2.5.7.11.23$	Mathieu
M_{24}	$2^{10}.3^3.5.7.11.23$	Mathieu
J_2	$2^7.3^3.5^2.7$	Hall, Janko
Suz	$2^{13}.3^7.5^2.7.11.13$	Suzuki
HS	$2^9.3^2.5^3.7.11$	Higman, Sims
McL	$2^7.3^6.5^3.7.11$	McLaughlin
Co_3	$2^{10}.3^7.5^3.7.11.23$	Conway
Co_2	$2^{18}.3^6.5^3.7.11.23$	Conway
Co_1	$2^{21}.3^9.5^4.7^2.11.13.23$	Conway, Leech
He	$2^{10}.3^3.5^2.7^3.17$	Held/Higman, McKay
Fi_{22}	$2^{17}.3^9.5^2.7.11.13$	Fischer
Fi_{23}	$2^{18}.3^{13}.5^2.7.11.13.17.23$	Fischer
Fi_{24}	$2^{21}.3^{16}.5^2.7^3.11.13.17.23.29$	Fischer
HN	$2^{14}.3^6.5^6.7.11.19$	Harada, Norton/Smith
Th	$2^{15}.3^{10}.5^3.7^2.13.19.31$	Thompson/Smith
B	$2^{41}.3^{13}.5^6.7^2.11.13.17.19.23.31.47$	Fischer/Sims, Leon
M	$2^{46}.3^{20}.5^9.7^6.11^2.13^3$	Fischer, Griess
J_1	$2^3.3.5.7.11.19$	Janko
$O'N$	$2^9.3^4.7^3.5.11.19.31$	O'Nan/Sims
J_3	$2^7.3^3.5.17.19$	Janko/Higman, McKay
Ly	$2^8.3^7.5^6.7.11.31.37.67$	Lyons/Sims
Ru	$2^{14}.3^3.5^3.7.13.29$	Rudvalis/Conway, Wales
J_4	$2^{21}.3^3.5.7.11^3.23.29.31.37.43$	Janko/Norton, Parker, Benson, Conway, Thackray

FIGURE 4

qui ne dépend en aucune manière de notre cerveau. Lorsque le mathématicien élabore des règles de cohérence logique, des règles d'exclusion, un formalisme, il construit un langage universel et celui-ci lui permet de reconnaître les propriétés de l'objet qu'il a au préalable construit. Finalement il ne « découvre » que les conséquences de ce qu'il a imaginé ! Il en révèle ce que Granger appelle le « contenu formel »¹. Personne n'ira dire, sauf peut-être certains croyants, que le Verbe existe avant la Matière !

5. Les mathématiques ne sont-elles qu'un langage ?

JPC : Quand nous parlons, nous manipulons des concepts. Tu décris une série de raisonnements, c'est-à-dire de procédures mentales ou cérébrales qui opèrent sur des objets concrets que tu te représentes. On peut penser au géomètre grec qui dessinait sur le sable des figures simples et étudiait leurs propriétés. Rien de ce que tu dis ne me persuade de la réalité de ces objets en dehors de notre cerveau. Même si tu arrives à préciser leur nombre ou leur nature, d'une manière parfaitement cohérente et organisée. Ce que tu dis tend au contraire à retirer aux objets mathématiques toute « réalité », au sens platonicien du terme. Tu conviens que les mathématiques constituent un langage, et qu'il existe plusieurs langages élémentaires... Peut-être les mathématiques constituent-elles la synthèse épurée de tous ces langages, une sorte de langage universel... Personne n'imagine que le chinois ou le russe ont existé avant l'homme

1. G. G. Granger 1988, *Sur l'idée de concept mathématique naturel*, Revue int. philosophie 167, p. 474-498.

dans l'univers. Alors, pourquoi cette hypothèse avec les mathématiques?

AC : « Rien ne prouve, dis-tu, la réalité de ces objets en dehors de notre cerveau. » Comparons la réalité mathématique au monde matériel qui nous entoure. Qu'est-ce qui prouve la réalité de ce monde matériel en dehors de la perception que notre cerveau en a? Principalement, la cohérence de nos perceptions, et leur permanence. Plus précisément, la cohérence du toucher et de la vue pour un seul et même individu. Et la cohérence entre la perception de plusieurs individus. La réalité mathématique est de la même nature. Un calcul effectué de plusieurs manières différentes donne le même résultat, qu'il soit fait par un seul individu ou par plusieurs. La vérité du théorème d'Euclide sur les nombres premiers ne dépend pas de tel ou tel mode de perception. Il est vrai que les mathématiques sont utilisées comme un langage par d'autres sciences. Mais on ne saurait, sans commettre une erreur grave, les réduire à n'être qu'un langage. C'est pourquoi la comparaison avec le chinois ne me paraît pas justifiée. On a commencé à explorer la réalité mathématique dans des zones où l'imagerie mentale liée au monde réel est très simple. C'est le cas pour la géométrie euclidienne. Ensuite, grâce aux procédés axiomatiques ou aux problèmes concrets posés par la théorie des nombres, on a pu accéder à des régions beaucoup plus éloignées de la réalité matérielle. Il n'empêche que la réalité à laquelle on est alors confronté est tout aussi solide que la réalité quotidienne. La frustration éprouvée par un mathématicien qui ne parvient pas à voir ce qui se passe dans cette réalité est tout à fait comparable à celle d'un aveugle qui cherche son chemin. Cela me suggère l'allégorie suivante : imagine que j'habite un village que je ne puisse quitter et qu'à une dizaine de kilomètres, se dresse une immense tour. Si j'étais le seul aveugle du village, mes voisins passeraient beaucoup de temps à me décrire cette tour dont l'existence ne ferait

pour eux aucun doute, tandis que je passerais le mien à leur expliquer qu'elle n'est qu'une construction mentale destinée à rendre compte de certains phénomènes visuels dont je n'ai que faire. Ainsi, malheureusement, tant qu'on n'est pas confronté à la réalité mathématique, on peut sans risque nier son existence.

JPC : Cette « cohérence de la perception » du monde extérieur, elle est due à ton appareil cérébral, mais à un *niveau d'abstraction* inférieur à celui des objets mathématiques. Qu'on puisse reconnaître aux objets mathématiques des propriétés universelles ne prouve pas plus leur indépendance par rapport au cerveau humain que l'existence du mot « Etat » ou du mot « bonheur ». A cette différence près que les concepts mathématiques ont une définition plus précise et plus restrictive, et de ce fait, possèdent des propriétés mieux définies, plus « universelles ».

D'autre part, il me semble que tu fais un usage répété de la métaphore. Tu compares la recherche mathématique à l'exploration d'un continent ou d'un village avec ses rues et sa tour. Mais cette métaphore fait chuter le discours d'un niveau mathématique abstrait à un niveau inférieur, concret et imagé, qui ne doit en aucune manière être pris au premier degré. Une métaphore ne saurait avoir valeur démonstrative. Pire, tu joues sur les sens multiples *et* contradictoires des mots « réalisme » ou « réalité ». Le « réalisme » c'est d'abord la doctrine platonicienne selon laquelle les *Idées* font partie d'un monde distinct du monde matériel, et ont une existence effective à un plus haut degré que les êtres individuels et sensibles qui ne seraient que leur reflet et leur image (*voir figure 1 ci-dessus*). Mais c'est aussi la doctrine d'après laquelle l'être est indépendant de la connaissance actuelle qu'en ont les sujets conscients. Enfin est « réaliste » celui qui postule une différence de nature entre l'être et la pensée : l'être ne peut ni être déduit de la pensée, ni s'exprimer de façon adéquate et exhaustive en termes

logiques. Hélas, tes métaphores te font passer du premier au troisième sens, alors que ces sens sont contradictoires ! J'utilise, quant à moi, le mot « réalisme », ou le terme « réalité », principalement en un sens non platonicien, qui est une sorte de compromis entre les deux autres définitions. Pour moi, la matière sous ses divers états, les êtres vivants, l'homme, existent indépendamment de la pensée humaine et de la connaissance *actuelle* que les sujets conscients en ont. Mais la pensée humaine, elle-même expression d'un état particulier de la matière, tente de décrire cet « en soi », cette *ultima actualitas*. Elle cherche à en donner, grâce à l'expérience, une définition progressive, mais pas nécessairement exhaustive. Je fais donc une distinction très nette entre la réalité de la matière et ce que tu appelles la « réalité mathématique ». L'existence de cette dernière me paraît liée à la pensée de l'homme, elle-même produit de l'Evolution des espèces.

II

Platon matérialiste?

1. *L'ascèse intellectuelle du matérialiste*

JEAN-PIERRE CHANGEUX : Tes thèses sur la nature des objets mathématiques paraissent un peu paradoxales : tu défends un point de vue platonicien tout en m'assurant du fondement matérialiste de ta pensée. Peut-être devrions-nous donc d'abord revenir sur ce qu'on peut appeler le matérialisme, ou plutôt la méthode, le *programme* matérialiste ? Il s'agit, comme le montre J.T. Desanti dans *La philosophie silencieuse*¹, d'une tentative d'explication qui requiert un matériau minimal, si possible limité aux lois physiques et chimiques. Le matérialisme suppose donc, pour employer le terme de Spinoza, une *emendatio intellectus*, une réforme de l'entendement, en forme d'« ascèse » intellectuelle, par laquelle on s'efforce d'éliminer les résidus mythiques qui nous hantent, en particulier le platonisme.

1. J.T. Desanti, *La philosophie silencieuse*, Seuil, Paris, 1968.

L'explication matérialiste contribue à réintégrer l'homme dans la nature. Dans cet ouvrage que je trouve excellent, Desanti montre que cette tâche suppose la construction de modèles de la réalité qui contiennent toujours, selon ses termes, un sous-modèle, à la constitution duquel on apporte le plus grand soin. Pour lui, « c'est le modèle de l'ensemble des procès qui produisent la connaissance et qu'il importe de construire de telle façon que : 1° il soit compatible avec le modèle de la réalité, 2° de lui soit éliminé explicitement tout appel à une forme quelconque de transcendance. Appelons, pour fixer les idées, *appareil de connaissance* le sous-modèle dont la constitution est ainsi exigée » ¹. Pour le neurobiologiste, l'appareil de connaissance qui permet de saisir la réalité, de construire ces modèles, c'est bien entendu le cerveau. Desanti, philosophe des mathématiques, pose clairement le problème de la nature des mathématiques en termes neurobiologiques, mais considère qu'il s'agit d'une utopie. Il va même jusqu'à écrire que « la constitution d'un modèle adéquat de l'appareil de connaissance ne peut être que chimérique », et qu'« il faut donc se résoudre à une épistémologie matérialiste faible » ². Abandon du philosophe, par une ignorance des neurosciences hélas fréquente, et, il faut le dire aussi, par insuffisance, au moment où Desanti écrit, des neurosciences et des sciences cognitives elles-mêmes.

Je défends au contraire une épistémologie matérialiste forte, la seule qui me paraisse acceptable de la part d'un scientifique averti, honnête avec lui-même. Ce programme n'est pas neuf. Il se trouve déjà formulé par Démocrite, ce philosophe présocratique qui avait, selon la légende, toujours le sourire (*voir figure 5*). Et nombreux ont été, au fil de l'histoire, les scientifiques qui ont eu le courage de l'adopter

1. *Ibid.*, p. 139

2. *Ibid.*, p. 145



FIGURE 5

Portrait de Démocrite par Antoine Coypel (1692).

Né vers 500-457, à Abdère, une des colonies ioniennes où les cultures grecque et orientale se trouvaient en contact, Démocrite aurait eu une vie extrêmement longue, de 100 à 109 ans. Contemporain de Socrate, il fut, avec Leucippe, le fondateur de l'atomisme, et selon Nietzsche, le premier penseur rationaliste à avoir éliminé tout élément mythique de sa pensée. Il est traditionnellement représenté avec un sourire qui témoignerait de sa joie d'avoir triomphé des craintes irrationnelles et des superstitions. (Musée du Louvre; d'après cliché couleur RMN.)

en dépit des persécutions dont ils furent l'objet : Vanini brûlé par l'Inquisition à Toulouse en 1619, l'anatomiste Vésale, et bien entendu Galilée... ne sont que quelques-unes de ces victimes d'une intolérance encore vive de nos jours.

Il faut donc définir les composantes de ce que Desanti appelle l'appareil de connaissance, et tenter d'en décrire les productions, en particulier mathématiques. L'appareil de connaissance est « un mécanisme d'abstraction ou de construction qui fabrique des types et des classes d'objets à partir d'un matériau sensible que le monde fournit en original ». C'est une excellente définition du fonctionnement cérébral. La tâche du neurobiologiste consiste donc, pour réaliser une épistémologie matérialiste forte, à décrire en particulier comment le cerveau de l'homme engendre les objets parmi lesquels se rangent entre autres les objets mathématiques. Que te suggère cette approche matérialiste ?

ALAIN CONNES : D'une part, il existe indépendamment de l'homme une réalité mathématique brute et immuable; d'autre part, nous ne la percevons que grâce à notre cerveau, au prix, comme disait Valéry, d'un mélange rare de concentration et de désir. Je dissocie donc la réalité mathématique de l'outil que nous avons pour l'explorer, et j'admets que le cerveau est un outil d'investigation matériel qui n'a rien de divin, qui ne doit rien à une transcendance quelconque. Mieux on comprendra son fonctionnement, mieux on pourra l'utiliser. Mais la réalité mathématique n'en sera pas pour autant changée. Pas plus que la suite des nombres premiers. Seule la somme de nos connaissances s'en trouvera modifiée. Si j'étais insensible au point de vue matérialiste, je pourrais facilement prétendre que « l'esprit humain » n'a que faire d'une meilleure connaissance du fonctionnement physique et biologique du cerveau. Loin de moi cette idée. Ma position est donc raisonnable.

JPC : Le mot « indépendant » nécessite une définition. Dans le cadre du réalisme platonicien, il signifie immaté-

rialité. Mais j'aimerais bien connaître le support de ces objets mathématiques dont tu prétends qu'ils existent indépendamment du cerveau de l'homme, tout en te déclarant matérialiste ! J'imagine difficilement que les nombres entiers existent dans la nature. Pourquoi ne pas voir $\pi = 3,1416$ écrit en lettres d'or dans le ciel, ou $6,02 \times 10^{23}$ apparaître dans les reflets d'une boule de cristal ! Les atomes existent dans la nature. Certes. Mais l'atome de Bohr, lui, n'existe pas. Une poule pourrait éventuellement évaluer le nombre d'œufs qu'elle a pondus, ou mieux, se rendre compte de l'espace qu'ils occupent dans le nid. Mais elle ne sait sûrement pas compter jusqu'à dix ni définir les propriétés des nombres entiers. Les mathématiques me paraissent plutôt constituer un langage formel, simplifié au maximum, et propre à l'espèce humaine.

AC : Je crois qu'il faut se garder de confondre la réalité mathématique et son illustration possible dans des phénomènes naturels. Quand je parle de l'existence indépendante de la réalité mathématique, je ne la localise absolument pas dans la réalité physique. Un certain nombre de modèles physiques utilisent, il est vrai, les mathématiques pour décrire des phénomènes naturels, mais ce serait une grave erreur de réduire les mathématiques à ces phénomènes. Je pense que le mathématicien développe un « sens », irréductible à la vue, à l'ouïe et au toucher, qui lui permet de percevoir une réalité tout aussi contraignante mais beaucoup plus stable que la réalité physique, car non localisée dans l'espace-temps. Lorsqu'il se déplace dans la géographie des mathématiques, le mathématicien perçoit peu à peu les contours et la structure incroyablement riche du monde mathématique. Il développe progressivement une sensibilité à la notion de simplicité qui lui donne accès à de nouvelles régions du paysage mathématique.

JPC : Ton argument principal est donc la simplicité. Mais ne peut-on pas dire également que, par exemple, le

thème de la Septième Symphonie de Beethoven est une mélodie d'une extrême simplicité?

AC : Oui, mais elle n'est pas nécessaire. Voilà la différence.

JPC : Mais, la nécessité, c'est ton cerveau qui la produit! C'est toi qui crées cette simplicité, lorsque tu confrontes tes représentations mentales entre elles ou à des objets naturels, lorsque tu constates leur adéquation ou leur inadéquation à l'aide du sens dont tu parles et que je considère comme produit de nos facultés cérébrales. Encore une fois, est-ce que cela prouve que cette simplicité a une origine immatérielle?

AC : La différence avec une symphonie de Beethoven est la suivante : en mathématiques, on peut démontrer, vraiment démontrer, une fois le problème bien posé, par exemple pour les corps finis, qu'on a obtenu la liste complète des objets qu'on cherche. Mais il n'y a aucun théorème qui permet de déduire d'un premier thème le reste de la symphonie de Beethoven.

JPC : Cette différence est importante. Mais cette propriété « générative » des mathématiques, on la retrouve sous une autre forme dans l'écriture musicale, en particulier chez Bach, chez Boulez et d'autres compositeurs contemporains. Elle constitue un des traits caractéristiques du langage humain dont l'expression la plus simple est la syntaxe. Les concepts eux-mêmes peuvent d'ailleurs posséder une certaine générativité. Considérons par exemple le concept de liberté. Bien que non mathématique, ce concept a eu, au moment de la Révolution française, un pouvoir génératif considérable qu'il conserve encore aujourd'hui. Que de concepts nouveaux, que de textes de lois se fondent sur sa définition! Toute une série de réorganisations sociales, de droits nouveaux et de bouleversements dans la structure de l'Etat en sont le résultat (*voir figure 6*). Pour autant, personne ne dira que la « liberté » existe, dans la nature, indé-



FIGURE 6

La statue de la Liberté, place Tian Anmen. Pékin, 29-30 mai 1989 (Copyright Agence Vu : Manuel Vimenet).

pendamment de l'homme. Bien entendu, la démonstration mathématique que tu offres est beaucoup plus rigoureuse, fermée, complète, cohérente, que sais-je, que ce que l'histoire a pu tirer du concept de « liberté ». Mais ne pourrait-on pas comparer un concept aussi abstrait que celui de liberté à celui de nombre entier, en dépit bien entendu des conséquences que chacun entraîne? Pourquoi poser une différence de *nature* aussi profonde?

AC : Ne confondons pas outil et réalité étudiée. Le concept de liberté a été élaboré progressivement par l'esprit humain pour rendre compte de certains comportements des êtres vivants. Je ne mets pas en doute leur réalité! De même, le mathématicien élabore des outils, comme la méthode axiomatique, ou des concepts, comme la topologie générale ou les probabilités, qui lui permettent par exemple de mieux comprendre la suite des nombres premiers. Mais qu'il y ait élaboration progressive des concepts et des méthodes d'investigation n'altère en rien la réalité de cette suite. Elle nous permet simplement de mieux la comprendre. Ta résistance à admettre l'existence de la réalité mathématique provient donc d'une part d'une confusion entre outils conceptuels et réalité, et d'autre part de l'existence d'une illustration physique très partielle des mathématiques.

JPC : Je ne confonds en aucune manière outils conceptuels et réalité, au sens où j'emploie ce mot. Puisque, pour moi, ces « outils » servent à étudier des *propriétés* d'objets que le cerveau du mathématicien produit et qui ont une authentique réalité physique. Par contre, je ne considère pas que la *méthode* axiomatique soit un concept. C'est une *procédure* cérébrale. Alors que le nombre entier est un concept, une « représentation mentale » simplifiée, dont on reconnaît facilement les propriétés originales. D'après moi, la « liberté » est un concept authentique et ne se compare pas à la méthode axiomatique.

2. *Psychanalyse des mathématiques*

AC : Un des traits essentiels du travail du mathématicien est de reconnaître la cohérence interne et le caractère génératif propre à certains concepts. Des concepts très simples arrivent à engendrer toutes sortes d'autres idées ou d'autres modèles. De proche en proche, on a vraiment l'impression d'explorer un monde... et d'atteindre une cohérence qui montre qu'on en a exploré entièrement une région. Dans ces conditions, comment ne pas sentir que ce monde a une existence indépendante ?

JPC : Tu dis « sentir » ? Ton attitude vis-à-vis des mathématiques serait-elle un sentiment plutôt qu'une réflexion ?

AC : C'est plutôt une intuition, une intuition laborieusement construite. Ma position se fonde d'une part sur la frustration que j'éprouve souvent face à des solutions partielles et contradictoires d'un problème, et d'autre part sur un contact direct avec les objets mathématiques, contact qui suscite une intuition évidemment disjointe de celle des phénomènes naturels. Le réalisme et le matérialisme ne me paraissent pas du tout inconciliables. Quel est le prix à payer pour accepter, comme hypothèse de travail, l'existence indépendante de la réalité mathématique ? Aucun, me semble-t-il. A l'inverse cela nous donne l'assurance qu'il y aura toujours moyen de communiquer ces concepts d'une civilisation à l'autre.

JPC : Le véritable prix à payer, c'est de comprendre comment notre « appareil de connaissance » produit des objets de cette nature ! Je me demande d'ailleurs dans quelle mesure l'indépendance dont tu parles ne vient pas simplement du fait qu'il s'agit d'objets culturels très particuliers, trans-

missibles d'un individu à l'autre quelle que soit leur culture, une sorte de sémantique universelle limitée, jusqu'à mieux informé, et en toute objectivité, à l'univers de l'homme. Que ces objets puissent exister sous forme écrite, par exemple tracés dans le sable comme le faisaient les premiers Grecs, ou enregistrés, comme aujourd'hui, sur les bandes magnétiques d'un ordinateur, laisse croire qu'il s'agit d'objets indépendants du cerveau humain. Or il n'en est rien. Ce sont plutôt des « représentations culturelles » susceptibles de se propager, de fructifier, et de proliférer et d'être transmises de cerveau à cerveau. Elles ont des propriétés spécifiques, en particulier cette cohérence, cette « nécessité intérieure », que tu te plais à souligner, et qui leur confère une « apparence » d'autonomie. C'est cette « apparence » qui te fascine, et la fascination que tu ressens est celle qu'exerce, après coup, l'objet créé aux yeux du créateur. Elle s'explique par la pratique scientifique elle-même, et tout ce qu'elle implique de *subjectivité*. Peut-on dire qu'un sujet en psychanalyse, par l'expérience qu'il acquiert de lui-même ou des personnes avec lesquelles il entretient des relations, progresse dans la compréhension de la nature profonde de son propre cerveau ? Hélas non. La psychanalyse n'a pas conduit à des progrès significatifs dans la connaissance du cerveau, de son architecture, de sa nature physico-chimique. Je crains que le « sentiment » que tu as de « découvrir » cette « réalité » toute platonicienne ne soit qu'une vision purement introspective, et de ce fait subjective, du problème. J'admets toutefois que les mathématiques constituent une production cérébrale d'une nature particulière. Et je pense qu'on pourrait s'entendre sur une définition de ce genre. Les objets mathématiques sont des concepts abstraits au même titre que le concept de liberté. Ils ont des propriétés spécifiques. Mais celles-ci n'impliquent en aucun cas leur immatérialité, non plus que le réalisme platonicien.

AC : Notre discussion tourne autour de la définition du

mot « réalité ». Pour moi, la réalité se définit par la coïncidence et la permanence des perceptions, soit d'un même individu, soit de plusieurs individus à l'intérieur d'un groupe.

JPC : Cette perception collective est nécessaire. Mais pas suffisante. Elle inclut tant les illusions optiques que les hallucinations collectives... Les Indiens Huichols, lors du pèlerinage annuel pendant lequel ils consomment des champignons hallucinogènes, ont *tous* le sentiment d'être *réellement* montés au Paradis. La « coïncidence des perceptions » ne suffit donc pas pour définir une réalité objective !

3. *Les objets mathématiques sont-ils des représentations culturelles comme les autres ?*

JPC : Mais ne peut-on dire qu'un objet mental se définit par sa cohérence interne, par un certain nombre de propriétés *exclusives* de toute autre et par le fait que plusieurs individus au sein d'un groupe, fort heureusement, sont capables de le percevoir en commun ? Cela n'a rien d'extraordinaire. Ils ont tous le même cerveau, ou presque ! D'autre part, j'ai déjà souligné que les mathématiques ont une histoire. Si ces objets mathématiques existaient dans l'univers de manière intemporelle, comme l'imaginent Pythagore et Platon, on devrait pouvoir les rencontrer à tout moment. Or les mathématiques évoluent, tant dans leur contenu que dans l'écriture et la symbolique. Pourquoi ce constant renouvellement que tu as évoqué ? On voit mal les objets mathématiques d'une *mathesis universalis* être remis en cause par une nouvelle théorie. S'ils étaient si universels et si indépendants de notre propre cerveau, pourquoi évolueraient-ils ? L'histoire des mathématiques n'est pas du tout linéaire. Elle est faite de controverses, de débats, de diver-

gences, de renouvellements, de remises à jour incessantes... Bref, on a l'impression d'avoir affaire à des objets culturels, qui sont produits et utilisés à chaque stade du développement de notre civilisation, et qui sont renouvelés au fur et à mesure que d'autres objets culturels, qui ne sont pas nécessairement mathématiques, évoluent eux-mêmes.

AC : Les connaissances mathématiques ont évidemment un caractère historique, tout comme l'exploration d'un continent. Devant la liste des mathématiciens qui, au prix d'efforts héroïques, ont découvert les groupes finis simples sporadiques, le profane n'a-t-il pas la même impression que devant une liste d'explorateurs? Pour reprendre mon exemple, la démonstration de la classification des groupes finis est pour le moment trop longue pour qu'un non-spécialiste puisse à lui seul la vérifier complètement avec une certitude irrévocable. Ce domaine fait donc partie d'une frange des résultats mathématiques qui ne sont pas encore stabilisés. A l'opposé, la liste des corps finis est relativement aisée à comprendre, et il est facile de démontrer qu'elle est complète. Elle s'insère dans la réalité mathématique complètement explorée, sur laquelle peu de problèmes subsistent. Il est bien évident que, dans la frange des recherches actuelles, le phénomène culturel et social contribue à indiquer les directions dans lesquelles on doit aller. Pour reprendre ma comparaison, la conquête du pôle Nord a certainement obéi, pendant un certain temps, au même type de motivations liées au contexte culturel et social. Mais une fois l'exploration achevée, ces phénomènes culturels et sociaux s'effacent, et seul reste un corpus parfaitement stable qui épouse le mieux possible la réalité mathématique et que nous nous efforçons d'enseigner aux générations futures. Cette vue un peu simplificatrice ne saurait nous empêcher de faire la distinction entre réalité mathématique établie et outil d'investigation.

JPC : Tu fais allusion à l'acquisition de connaissances.

Les connaissances que nous avons de l'univers en général sont de même nature. On ne va pas remettre en question le fait que la terre tourne autour du soleil !

AC : Une fois un théorème mathématique démontré, comme celui d'Euclide sur les nombres premiers, personne ne le remettra en question.

JPC : Je n'en ai jamais douté. Je pense que l'on peut s'entendre encore sur ce point. Mais ce qui me frappe surtout, c'est ton emploi de la notion de *frange* et plus précisément de « frange des résultats mathématiques *non stabilisés* ». Au départ ont été créés un petit nombre d'objets mathématiques relativement simples. Progressivement, cette frange s'est amplifiée. La « stabilisation » dont tu parles me paraît liée à l'environnement culturel. C'est pour cette raison que j'appelle les objets mathématiques des objets culturels. Au fil de l'histoire, seule une fraction des objets mathématiques produits par le cerveau des hommes de l'art a été retenue, *sélectionnée*, engrangée dans le cerveau de leurs collègues, puis dans des textes. Certains auteurs vont même jusqu'à dire que les mathématiques sont nées le jour où des philosophes grecs ont commencé à dessiner des figures dans le sable, où ils ont pu utiliser une mémoire autre que leur mémoire à court terme, qui ne permet pas de stocker tous ces objets. Donc, ce patrimoine culturel peut, année après année, avoir été simplement mis en forme, et réduit à une structure minimale cohérente, jusqu'à constituer ce que tu appelais un corpus. Ce dernier devrait ainsi son existence aux facultés cérébrales de l'homme qui lui ont permis d'établir une forme de dialogue entre, d'une part, sa mémoire à court terme, sa mémoire de travail, qui utilise les objets mathématiques, et d'autre part, une mémoire extérieure non cérébrale qui les emmagasine pour les rendre publics. L'homme a ainsi pu utiliser des objets mathématiques, extérieurs à son propre cerveau, pour en produire de nouveaux, les comparer aux précédents, les « passer au crible de la

raison », et les remettre dans le stock commun, une fois certain qu'ils s'intègrent à ce stock. L'évolution contingente qui me semble, de l'extérieur, dominer le champ des mathématiques, pourrait donc nous amener à définir les objets mathématiques comme des objets culturels, des représentations publiques d'objets mentaux d'un type particulier, qui se produisent dans le cerveau des mathématiciens et se propagent d'un cerveau à l'autre... jusqu'à celui des biologistes.

4. Le darwinisme des objets mathématiques

AC : Il est certain que l'exploration de la réalité mathématique subit des influences culturelles. Cela ne suffit pourtant pas à justifier l'appellation d'objet culturel. Je pense que la difficulté principale, c'est la distinction entre la réalité mathématique « crue » et les outils de pensée mis au point par les mathématiciens pour l'appréhender. Ces outils font effectivement partie de notre patrimoine culturel. Prenons l'exemple de ce que j'appellerai « l'étude des comportements asymptotiques ». Il se peut que la réalité mathématique soit trop complexe pour être facilement perceptible. Par exemple, il n'existe aucune formule simple qui donne le n ième nombre premier. Le problème asymptotique consiste à trouver une formule qui donne approximativement l'ordre de grandeur du n ième nombre premier. On arrive ainsi à démontrer que le nombre de nombres premiers plus petits qu'un entier n est équivalent au quotient de n par son logarithme. On dévoile ainsi un aspect de la réalité mathématique. Mais, j'insiste, il faut distinguer l'outil de pensée et la réalité mathématique

qu'il explore. Un mathématicien peut très bien *inventer* un nouvel outil de pensée. Tant qu'il n'a pas réussi, grâce à lui, à dévoiler une nouvelle partie de la réalité mathématique que ses contemporains ne connaissaient pas, ceux-ci considéreront l'outil créé avec une certaine incrédulité. Il ne suffit pas d'avoir de l'imagination pour faire des mathématiques !

JPC : Nous arrivons à un moment intéressant de la discussion. Tu as spontanément évoqué l'évolution des connaissances mathématiques. Je reprends cette idée parce qu'elle s'intègre à celle d'une évolution générale des connaissances, des objets culturels sous leurs diverses formes. Tout laisse à penser, tu en conviens, que les objets mathématiques, comme tout objet de connaissance, sont apparus par « mutation mentale », au hasard des expériences cérébrales des mathématiciens. Ils ont ensuite été utilisés, exploités, triturés si j'ose dire, par le raisonnement. Puis s'est déposé un « résidu » sélectionné – j'emploie intentionnellement le terme darwinien ¹ – en vertu de son adéquation à l'ensemble déjà existant, et pour des raisons de cohérence. De ce fait, une *rigidification* s'est produite. C'est sur ce point que je ne suis plus d'accord avec toi. Cette cohérence et cette rigidité me paraissent résulter *a posteriori* de l'évolution. Permetts-moi de comparer l'évolution des objets mathématiques à l'évolution biologique. Même s'il existe un apparent « progrès » continu dans l'évolution des vertébrés, des poissons aux amphibiens, des reptiles aux mammifères, puis des singes à l'homme (*voir figure 7*), personne actuellement ne va imaginer, à part quelques croyants comme Teilhard de Chardin, que l'évolution s'est déroulée avec, pour finalité, l'homme et sa perfection. Je n'irai pas jusqu'à comparer ton attitude à celle de Teilhard de Chardin, mais je crains,

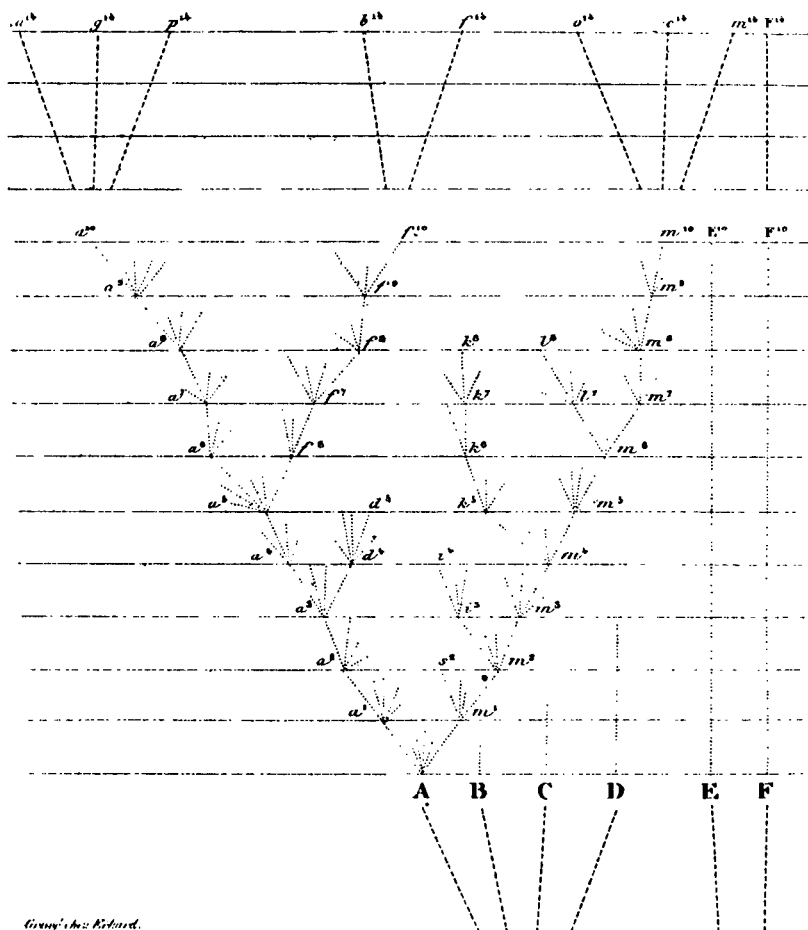
1. Voir J.-P. Changeux, *L'homme neuronal*, Fayard, Paris, 1983, et G. Edelman, *Neural Darwinism*, Basic Books, New York, 1987.

lorsque tu parles du mathématicien « dévoilant » progressivement un univers mathématique et structuré, d'y voir une sorte de finalisme que j'arrive à comprendre de la part du praticien, mais que je trouve inattendu de la part du théoricien.

Dans nos laboratoires de biologie, lorsque nous examinons une molécule, nous nous demandons comment il se fait qu'elle ait une activité enzymatique, qu'elle soit le support de l'hérédité... Nous nous posons la question de la « cause prochaine », comme la distingue le biologiste évolutionniste Ernst Mayr ¹. Nous nous demandons à quoi elle sert. Cela ne veut pas dire que cette molécule a été conçue, par un Etre tout-puissant, pour faire ceci ou cela, ni qu'elle s'intègre à un univers rationnel conçu par un Esprit infiniment intelligent. Ces métaphores issues d'Aristote se rencontrent en permanence dans le langage de laboratoire. J'imagine que les mathématiciens, comme tous les scientifiques, les utilisent dans leur pratique. Mais personne ne prend plus au sérieux, en tout cas en biologie, les thèses finalistes. Spinoza déjà, dans la rigueur de sa méthode philosophique, mettait en garde contre la tendance dangereuse du raisonnement humain à user d'arguments de type finaliste. Et je me demande dans quelle mesure cette cohérence et cette rigidité que tu évoques ne ressemblent pas à la cohérence des organes du mammifère et à la rigidité de son squelette. N'oublie pas qu'on a longtemps cru que l'Univers, et en particulier les êtres vivants, étaient des créations divines que le naturaliste « découvrait » par son travail, saisissant ainsi une harmonie préétablie du monde!

AC : Entendons-nous bien sur le terme « évolution ». En mathématiques, comme dans toute autre discipline, les

1. E. Mayr, How biology differs from the physical sciences, in D.J. Drepew et B.H. Weber, *Evolution at a cross road. The new biology and the philosophy of science*, MIT-Press, Cambridge, 1985, p. 45-63.



Gravé chez Richard.

FIGURE 7

Figure originale de Darwin, extraite de *L'Origine des Espèces*, qui illustre, selon ses termes, « l'effet probable de l'action de la sélection naturelle, par suite de la divergence des caractères et de l'extinction, sur les descendants d'un ancêtre commun ». (6^e édition anglaise, traduction E. Barbier, p. 125, C. Reinwald, Paris.)

connaissances évoluent. Mais la réalité sur laquelle elles portent ne change pas. Une fois établie et démontrée, la liste des groupe finis simples par exemple ne changera jamais. Elle est vraiment le produit d'une découverte.

Que vient alors faire ici le finalisme? Je ne pense pas qu'affirmer l'existence d'une réalité mathématique indépendante de sa perception soit une thèse finaliste. A aucun moment, je n'oserais affirmer que tel ou tel objet mathématique obéit à une finalité quelconque. Jamais un mathématicien n'emploierait un argument de cette nature!

5. *Les croyances en mathématiques*

AC : Je ne suis donc nullement finaliste. Et je ne pense pas pouvoir modifier ma position...

JPC : Et pourquoi pas?

AC : Non, non...

JPC : Mais peut-être t'es-tu lancé dans le débat avec des idées un peu arrêtées...

AC : J'y crois...

JPC : Attention, tu emploies le mot « croire »!

AC : Oui, bien sûr. Mais une partie de la discussion est métaphysique...

JPC : Et elle est fondamentale.

AC : Bien sûr, pour autant qu'elle nous conduit à préciser la notion de réalité. J'admets, par humilité, que le monde mathématique existe indépendamment de la manière dont nous le percevons, et n'est pas localisé dans le temps et dans l'espace. Mais notre manière de l'appréhender obéit, elle, à des règles très voisines de celles de la biologie. L'évolution de notre perception de la réalité mathématique développe en nous un sens nouveau qui nous permet d'accéder

à une réalité qui n'est ni visuelle, ni auditive, mais d'une autre nature.

JPC : A cet égard peut-être allons-nous nous rencontrer à nouveau. Lorsque tu dis que notre cerveau développe un *nouveau* sens, tu deviens progressivement constructiviste. Je n'exclus pas du tout, et je parle en neurobiologiste, que notre appareil cérébral possède une flexibilité et une capacité de réorganisation telles qu'il puisse appréhender des objets d'une forme nouvelle qu'il n'avait pas l'occasion de percevoir dans le monde où il s'est formé il y a quelques millions d'années dans les plaines de l'Afrique centrale. Cette aptitude devrait lui permettre de produire, de capturer « un sens nouveau ». Mais cela n'implique pas nécessairement qu'il existe dans la nature un système mathématique totalement organisé que nous découvririons progressivement. Je pense que ta position recèle une contradiction : tu admetts, d'un côté, que les mathématiques évoluent selon un modèle conforme à celui que défendent les biologistes, et de l'autre, que le corpus mathématique constitue une *mathesis universalis*, un immense ensemble cohérent et stable dont on ne connaîtrait que des rudiments. Cette discussion me rappelle à nouveau les réflexions d'Ernst Mayr au sujet de la causalité dans les sciences de la vie. Il oppose la « cause prochaine », le « comment se fait-il ? » du biologiste ou du physiologiste à la « cause dernière », au « pourquoi » du métaphysicien. Et sa réponse est claire. La science du « pourquoi » n'est pas Dieu mais la biologie évolutionniste. Le pourquoi de l'existence des mathématiques est l'*Évolution* tant de notre appareil de connaissance que des objets mathématiques eux-mêmes. Quand tu évoques un sens nouveau qui se forge, qui permet d'accéder à des objets culturels, à leur tour en évolution, je pense que nous tombons alors d'accord.

III

La nature habillée sur mesure

1. *Les mathématiques constructivistes*

JEAN-PIERRE CHANGEUX : Le débat entre le mathématicien créationniste que tu es et le biologiste évolutionniste que je suis s'est quelque peu radicalisé. Nous arrivons toutefois à nous entendre sur plusieurs points : la définition des objets mathématiques comme représentations culturelles d'un type très particulier et le fait que les connaissances mathématiques progressent. Nous divergeons pour l'instant, mais peut-être évoluerons-nous l'un ou l'autre au fil de ces entretiens, quant à l'existence d'une réalité mathématique qui préexiste dans l'Univers au cerveau du mathématicien. D'après toi, le mathématicien ne fait que découvrir pas à pas cette *mathesis universalis* à laquelle tu crois – j'emploie délibérément ce terme. Toutefois, ton attitude ne correspond pas à celle de tous les mathématiciens. Emmanuel Kant, au XVIII^e siècle, soutenait déjà que « l'ultime vérité des mathématiques se trouve dans la possibilité qu'a l'esprit humain

d'en construire les concepts ». Nombre de mathématiciens, qu'on appelle constructivistes, pensent qu'un objet mathématique n'existe que dans la mesure où on peut le construire. Le débat semble d'ailleurs presque aussi vif entre formalistes et constructivistes qu'entre toi et moi ¹. L'un d'eux, Allan Calder, va même jusqu'à écrire que « les critères d'acceptabilité des mathématiques constructives sont plus rigoureux que ceux des mathématiques non constructives » ², et qu'à prendre le problème sous l'angle de la constructivisation, on gagne « une meilleure analyse et des théorèmes plus puissants » ³. Il est quand même remarquable que certains mathématiciens défendent des thèses, différentes des tiennes, et proches de celles du neurobiologiste que je suis. Allan Calder est encore plus direct que je l'étais lorsque j'évoquais ton vécu de mathématicien créateur et la subjectivité de ton attitude. Il écrit : « La majorité des mathématiciens d'aujourd'hui, formés au moule du formalisme depuis plusieurs générations, se trouvent placés dans une situation de *blocage mental* qui leur permet difficilement d'avoir une vue objective des mathématiques, au point que certains considèrent même que le constructivisme serait un cancer qui détruirait les mathématiques ⁴. » Voilà beaucoup de passion, d'irrationalité même, dans le débat entre mathématiciens. Au point que Calder termine son article dans le ton où se concluait notre dernier entretien : « Croire en l'existence d'une vérité mathématique en dehors de l'esprit humain exige du mathématicien un acte de foi dont la plupart ne sont pas conscients ⁵. » Nous voilà bien loin de l'*emendatio intellectus* chère à Spinoza!

1. M. Kline, *Mathématiques : la fin de la certitude*, traduit de l'anglais par J.P. Chrétien-Goni et C. Lazzeri, C. Bourgois, Paris, 1989.

2. A. Calder, « Les mathématiques constructives », in *Les mathématiques aujourd'hui*, préface de M. Berger, Pour la science, 1986, p. 203-211.

3. *Ibid.*, p. 210.

4. *Ibid.*, p. 211.

5. *Ibid.*, p. 211.

ALAIN CONNES : La distinction entre constructivisme et formalisme est avant tout méthodologique. On peut comparer les constructivistes aux alpinistes qui attaquent les parois à mains nues et les formalistes à ceux qui s'autorisent à prendre un hélicoptère pour survoler le sommet. Les deux attitudes ont des avantages qui dépendent des problèmes considérés. Sans aller jusqu'au constructivisme, on a parfois besoin, même pour les mathématiques courantes, de tempérer les effets de l'axiomatique établie, en particulier ceux de l'axiome du choix non dénombrable. Prenons un problème précis que j'ai rencontré, à l'occasion duquel les deux points de vue sont très différents. Il s'agit d'un débat assez ancien sur le problème de la mesurabilité, au sens de Lebesgue, des fonctions à valeur réelle. On arrive à démontrer qu'en utilisant seulement l'axiome du choix dénombrable, on ne peut construire de fonctions non mesurables. Il en résulte que le raisonnement mathématique qui n'utilise que l'axiome du choix dénombrable ne se heurtera jamais au problème de non-mesurabilité. En théorie constructiviste, on n'utilise jamais l'axiome du choix, et on ne se heurte donc jamais au problème de non-mesurabilité. Voyons maintenant le point de vue formaliste. Lorsqu'on édifie la théorie des ensembles en se fondant sur l'axiome du choix non dénombrable, on démontre que tout ensemble peut être bien ordonné. Or, un bon ordre sur la droite réelle est par essence non mesurable et bien entendu non constructible. L'axiome du choix non dénombrable simplifie considérablement la théorie des cardinaux, et donne donc une vision d'une partie de la réalité mathématique que je qualifierai d'assez grossière. On considère en particulier comme isomorphes, comme ayant la même cardinalité, des ensembles pour lesquels on ne peut effectivement construire aucune bijection de l'un à l'autre. Par exemple, l'ensemble des quasi-cristaux de Penrose a, avec l'axiome du choix non dénombrable, la cardinalité du continu, alors qu'il est

impossible de construire une bijection effective entre cet ensemble et le continu. Pour prendre à nouveau une image simpliste, l'axiome du choix non dénombrable donne une image « vue d'avion », donc simplifiée, de la réalité mathématique. Il est vrai que la plupart des mathématiciens ont été formés au moule de la théorie des ensembles, qui suppose l'axiome du choix non dénombrable, et qu'ils ne se rendent pas compte de sa fonction simplificatrice. Mais cette simplification n'intéresse qu'une toute petite proportion des mathématiques courantes, et en général elle est la bienvenue. On voit donc que des éclairages différents révèlent des aspects différents de la réalité mathématique, et il n'y a là rien de contradictoire. Le constructivisme ne remet donc pas en question l'existence d'un monde mathématique indépendant...

JPC : C'est pourtant ce que disent ses défenseurs. On ne peut pas les accuser d'obscurantisme. Ils connaissent tout de même l'univers des mathématiques. Mais, pour eux, il n'existe que dans la mesure où ils peuvent le construire pas à pas.

AC : Je pense que tu ne trouveras pas un constructiviste qui n'accepte la liste que j'ai donnée plus haut des groupes finis simples. Il faut bien comprendre que la plupart des objets mathématiques fondamentaux ont un caractère constructible qui explique que leur existence n'est pas remise en question par les constructivistes. Par contre, dans la méthode, la différence peut être considérable. Prenons par exemple un outil de démonstration très utile qu'on appelle les ultraproducts. Si le résultat à démontrer n'utilise pas, dans sa formulation, les ultraproducts, on sait, ce qui est un résultat de logique mathématique, qu'il existe une démonstration qui n'utilise pas cet outil. Il n'empêche qu'en général, pour certains problèmes, il est beaucoup plus facile de trouver la démonstration en utilisant les ultraproducts. On montre par exemple que les corps obtenus en faisant

des ultraproducts de corps péadiques sont les mêmes que ceux obtenus en faisant des ultraproducts de corps de séries formelles sur les corps finis F_p . Pour résoudre certaines équations, cela peut être très utile. Dans ce cas précis, on voit bien que c'est le point de vue constructiviste qui, en interdisant l'utilisation des ultraproducts, est conservateur et limitatif.

Quoi qu'il en soit, je ne crois pas que soit remise en cause l'existence d'un monde mathématique indépendant de nous et échappant à nos sens.

JPC : C'est ce que tu crois. La distinction constructivisme/intuitionnisme serait, selon toi, plus méthodologique qu'ontologique. Ce n'est pas ce que pensent les constructivistes... Ils prétendent vous remettre en question. Mais après tout, ton expérience subjective de mathématicien et ton acte de foi si véhément, puisque tu admetts que tu en fais un, sont peut-être révélateurs d'une vérité plus profonde qui ne ferait d'ailleurs appel à rien d'immatériel. Une vérité que tu ressens, que tu perçois, que tu imagines, mais sur laquelle nous n'arrivons pas à nous entendre, faute, peut-être, d'une conception d'ensemble qui puisse nous réunir.

Cette question de l'existence d'un monde mathématique constitue notre point majeur de désaccord. Entrant dans ton jeu, j'ai essayé de voir où ce monde pouvait se situer, quelle en était la *trace* dans la nature. Si tu formes l'hypothèse que ce monde mathématique existe en dehors de nous et si tu te dis matérialiste, tu es conduit à lui donner une base matérielle. Je ne vois pas sous quelle autre forme que l'organisation même de la matière, ce monde mathématique pourrait être présent dans la nature, en dehors, bien entendu, de ce qui est stocké dans les livres et dans la mémoire du cerveau des mathématiciens. Il existe incontestablement des régularités dans la nature qui nous entoure : le mouvement des planètes (*voir figure 8*), l'organisation des atomes dans un cristal de sel gemme, ou

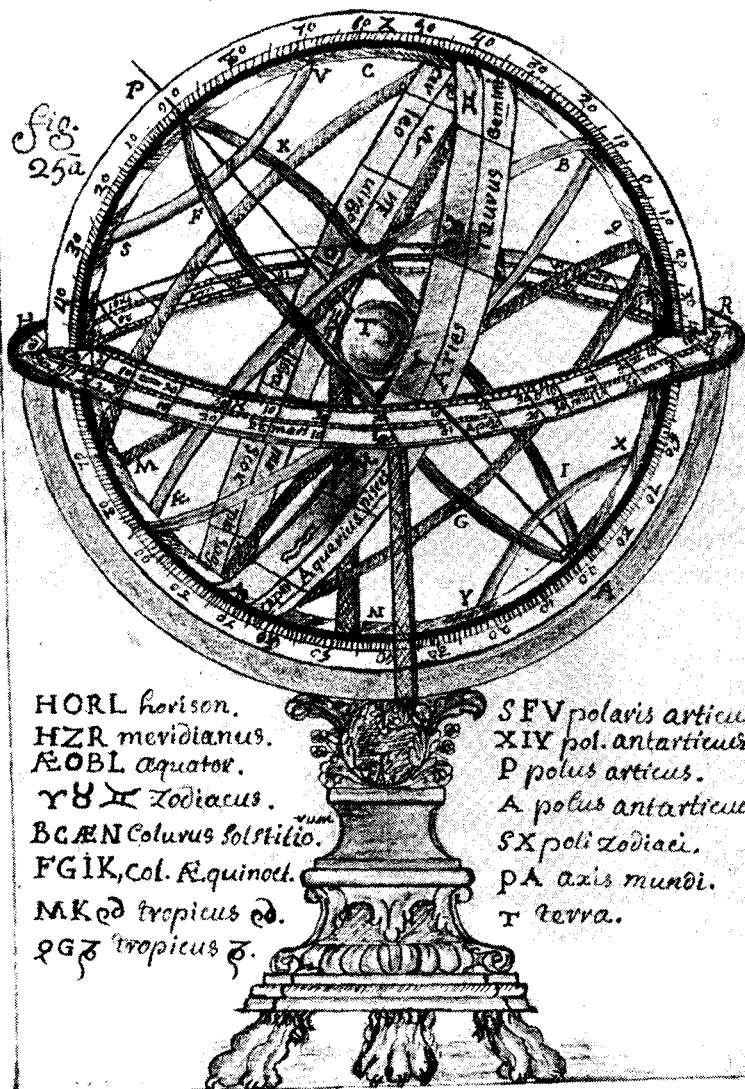
l'organisation en double hélice de l'acide désoxyribonucléique. Crois-tu que ces régularités sont l'expression d'une mathématique universelle qui constituerait en quelque sorte le « squelette idéal » de la matière autour de laquelle elle s'organiserait ? Ou bien penses-tu au contraire que ces régularités représentent, comme je l'accepte, des propriétés intrinsèques de la matière, mais qui ne sont pas nécessairement l'expression d'une loi mathématique originelle ? S'il en était ainsi, la tâche du scientifique « naturaliste » consisterait alors à appréhender ces régularités, à forger des outils, et à créer un langage et des concepts, en général mathématiques, qui permettent de les décrire. Pour choisir entre ces deux points de vue, il faut confronter ces régularités extérieures avec les objets mathématiques. Si les mathématiques étaient le principe organisateur de la matière, on devrait tôt ou tard trouver une adéquation *parfaite* entre régularité des objets matériels et régularité des objets mathématiques. Sinon, les mathématiques, produits du cerveau humain, ne sont qu'un langage approximatif qui nous sert à décrire une matière qui, en grande partie, nous échappe.

Les biologistes, comme les physiciens, dans leur démarche hypothético-déductive, construisent des objets de pensée ou modèles, qu'ils confrontent au réel physique qui leur est extérieur. Ces modèles sont des représentations simplifiées d'un objet ou d'un processus qui sont cohérentes, non contradictoires, minimales et validables par l'expérience. Un bon modèle est prédictif, en ce sens qu'il doit

FIGURE 8

Sphère armillaire. Les sphères armillaires sont des objets de métal ou de bois composés de cercles emboîtés, articulés entre eux, et censés reproduire le mouvement des planètes et de la sphère céleste autour de la terre (T), en fonction des mois de l'année. Construits principalement au cours du XVIII^e siècle, ils constituent, en quelque sorte, les premiers modèles mécaniques de l'univers. Le dessin de sphère armillaire présenté ici provient d'un cahier de cours manuscrit de l'étudiant G.L., rédigé en 1713. (Collection particulière.)

Sphaera Armillaris.

fig.
25^a

HORL horizon.

HZR meridians.

EOBL aequator.

YX Zodiaccus.

BCEN Colurus solstilio.

FGIK, col. Aequinoct.

MK ed tropicus ed.

QG ed tropicus q.

SFV polaris arcticus

XIV pol. antarcticus.

P polus arcticus.

A polus antarcticus.

SX poli zodiaci.

PA axis mundi.

T terra.

conduire à des expériences qui enrichissent nos connaissances. Il est également génératif puisqu'il peut susciter d'autres modèles théoriques et enrichir ainsi la théorie. Enfin, à la différence des « croyances » qui caractérisent une tradition culturelle définie, un modèle est *révisable*. Il me semble en effet nécessaire d'admettre que la plupart des modèles que la science produit sont valides à un moment donné de l'histoire de la science et qu'un certain nombre au moins de leurs propositions peuvent être révisées ou amendées. Tout cela, bien entendu, permet un progrès cumulatif de la connaissance. Nous utilisons ces objets de pensée pour capturer les régularités du monde physique, pour les décrire sous forme mathématique, une fois « sélectionné » le modèle qui paraît le plus adéquat. Dans notre démarche, nous abordons les régularités naturelles d'une manière indirecte. Nous essayons, en quelque sorte, de les *habiller* d'un certain nombre d'objets de pensée, dont, entre autres, les objets mathématiques. Mais cela n'implique pas nécessairement l'identification de ces objets naturels aux mathématiques que nous utilisons pour les décrire.

2. *L'efficacité surprenante des mathématiques*

AC : Loin de moi l'idée que la réalité mathématique se situe dans le monde physique. Je ne cherche donc pas du tout à identifier certains objets naturels aux mathématiques. Une fois la réalité mathématique distinguée de la réalité physique se pose pourtant le problème de leurs relations. Je commencerai par donner un exemple de ce qu'Eugene Wigner appelle « l'efficacité déraisonnable des mathématiques », qui, en général, ne résulte justement pas de l'effort pour habiller adéquatement les régularités naturelles. C'est

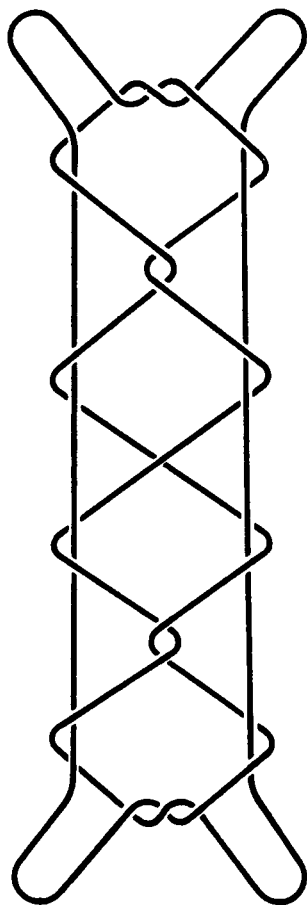
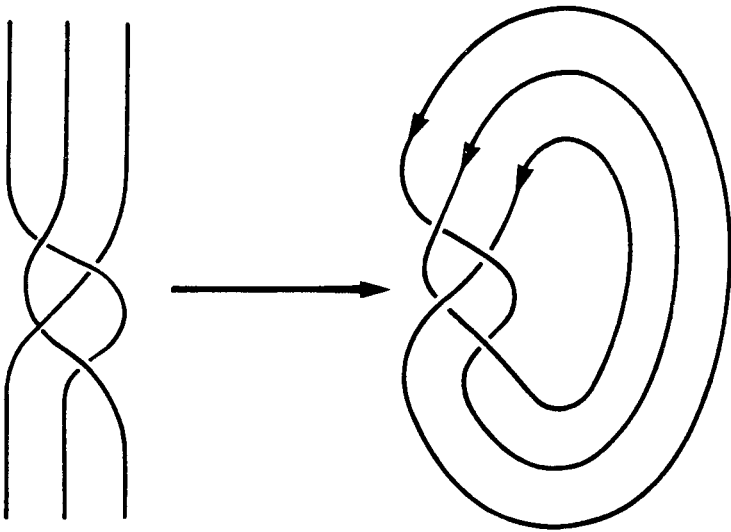


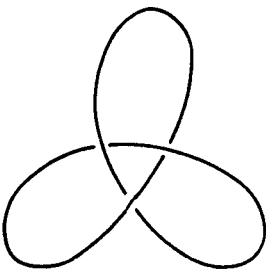
FIGURE 9

L'échelle de Jacob. Elle est un nœud connu de certains enfants, qui le réalisent en faisant passer une boucle de corde entre quatre doigts et en attrapant des brins libres entre les autres doigts. Cette échelle de Jacob est évidemment équivalente au nœud trivial en boucle. Esquimaux et Indiens d'Amérique du Nord sont friands de ces jeux de corde qui illustrent les infinies possibilités de motifs géométriques réalisables à partir d'une boucle, le plus simple des nœuds. (La théorie des nœuds.)

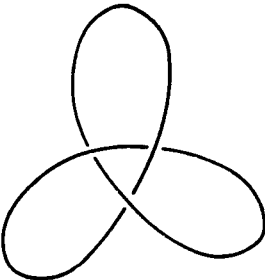


Tresse

Nœud associé



Nœud de trèfle



Sa réflexion

FIGURE 10

celui de la théorie des nœuds ¹ (*voir figures 9 et 10*). Quand on prend une ficelle et qu'on fait un nœud assez compliqué, le problème se pose de savoir si ce nœud peut être dénoué sans qu'on ait besoin de recourir à la méthode du nœud gordien. De fait, une magnifique théorie mathématique permet dans beaucoup de cas de résoudre ce problème : on l'appelle la théorie des nœuds. Un progrès très important dans cette théorie a été récemment effectué par un mathématicien dont la motivation de départ n'avait rien à voir avec les nœuds. Le Néo-zélandais V. Jones avait commencé par travailler avec moi sur un autre sujet. Ensuite il s'était intéressé à un problème très délicat d'analyse en dimension infinie. Il s'agissait de classifier les sous-facteurs d'un facteur donné, notion on ne peut plus éloignée de la théorie des nœuds. Il a travaillé seul pendant longtemps, et personne ne croyait à l'intérêt de ce qu'il faisait. On se demandait vraiment pourquoi il perdait ainsi son temps. Quelques années plus tard, il a réussi à démontrer que l'indice du sous-facteur prenait des valeurs discrètes ou un spectre continu. Il a découvert qu'au cours de sa démonstration apparaissait un groupe connu sous le nom de « groupe des tresses », dont on peut se faire une idée très concrète à partir de tresses, simplement en regardant des brins qui sont ou ne sont pas noués. C'était d'abord une simple image. Quand il faisait une conférence, il illustrait la présentation de ce groupe en dessinant des tresses. C'est alors qu'à New York il a rencontré une topologue, Birman, et qu'en discutant avec elle, il s'est aperçu que sa construction, la trace qu'il avait construite sur l'algèbre du groupe de tresses, donnait en fait un nouvel invariant pour les nœuds. Il a alors calculé cet invariant sur les nœuds les plus simples, ceux qu'on appelle les nœuds de trèfle, et il s'est aperçu

1. « La théorie des nœuds », par E. Belaga, in *Les progrès des mathématiques*, Pour la Science, Paris.

que, lorsqu'on prenait l'image du nœud de trèfle dans un miroir, son invariant n'était pas le même. Ce fut une surprise, car les invariants classiques sont invariants par réflexion. Ensuite, pour de très nombreux nœuds, il a fait toutes sortes de calculs nouveaux de son invariant, qui se calcule de manière très simple, et dont on ne connaît toujours pas d'interprétation purement géométrique. C'est un invariant extrêmement puissant qui permet de distinguer des nœuds que personne ne pouvait distinguer auparavant. Il permet par exemple de contrôler un nombre qu'on appelle le nombre gordien et qui a un sens évident. On essaye de faire traverser un brin à travers un autre jusqu'à ce que le nœud soit dénoué, et on compte le nombre de brins qu'il faut faire traverser pour arriver à dénouer le nœud. Eh bien, son invariant permet de mesurer le nombre gordien ! Et c'est tout à fait extraordinaire, parce que Jones est parti d'un problème de mathématique entièrement pure, caché dans les coins les plus obscurs, dans les régions les plus désertiques de la géographie des mathématiques. Mais la solution de son problème l'a conduit directement aux nœuds, qui, comme tu le sais, sont utiles en biologie, puisqu'ils entrent dans toutes sortes de problèmes de codages de molécules très compliquées comme les polymères. Son souci actuel est d'ailleurs l'application de son résultat à des problèmes très concrets. Voilà qui illustre parfaitement la puissance difficilement explicable des mathématiques quand on les pratique pour elles-mêmes, sans idée préconçue d'éventuelles applications.

JPC : L'histoire que tu décris comporte quand même une expérience.

AC : Ce n'était pas une expérience. C'était plutôt une coïncidence.

JPC : Oui, enfin... quand il a rencontré cet autre topologue et que cette rencontre a conduit à mettre en relation certains types d'outils mathématiques avec un problème plus

concret. Les nœuds peuvent exister dans la nature, et sont, le plus souvent, le produit d'une activité créatrice de l'homme. Mais je n'arrive pas à concevoir que la *théorie* des nœuds, elle, existe dans la nature *avant* que toute cette collection des diverses formes de nœuds connus ait été rassemblée. Cet excellent mathématicien a tout simplement réussi à créer un outil de pensée, que tu appelles un *invariant*, et il s'en est servi... Il a forgé un outil, comme l'homme a construit la roue pour avancer plus vite sur le sol. Au lieu d'enchaîner des dizaines de raisonnements, il a créé un « condensé de rationalité » qui lui a permis de résoudre instantanément le problème.

AC : Ce qui me frappe surtout, c'est que sa recherche, et sa découverte n'étaient pas du tout motivées par le problème des nœuds. C'est un exemple très intéressant de découverte motivée par des problèmes profonds de mathématiques pures. Sa recherche sur les facteurs l'a conduit à découvrir une fonction centrale sur le groupe de tresses. Tant que cette fonction lui était utile uniquement pour son problème de classification des sous-facteurs, elle n'avait évidemment aucun lien apparent avec les nœuds. Mais il péchait simplement par ignorance. Sa rencontre avec Birman lui a appris qu'en théorie des nœuds on utilise aussi le groupe des tresses et qu'on cherche, à cause d'un théorème de Markov, une fonction sur le groupe de tresses qui vérifie telle et telle propriété. Et il s'est écrié : « Mais je l'ai, je l'ai dans ma poche. »

JPC : Je comprends ce que tu veux dire. Deux démarches, au départ totalement indépendantes, en sont venues à converger. L'objet mathématique créé par l'un a ouvert la serrure qui restait fermée pour l'autre. Cela ne signifie pourtant pas que préexistaient une clé et une serrure prêtes à être découvertes !

AC : Je ne sais pas.

JPC : On touche ici le problème de fond de savoir pour-

quoi certains outils mathématiques, créés indépendamment de toute investigation sur les particules, sur les nœuds et sur les autres objets naturels, ont une telle faculté d'adéquation...

AC : Tout à fait. C'est ce qu'on appelle l'efficacité déraisonnable des mathématiques.

JPC : Je voudrais quand même que tu me dises jusqu'où elle va, cette efficacité, quelle est son universalité? J'ai noté chez les physiciens et chez certains mathématiciens une tendance à s'enticher d'un modèle mathématique à la mode. Ils le considèrent comme bon à tout faire et le plaquent sur des collectivités d'atomes, de neurones, de fourmis, d'hommes. Tu as une expérience vécue des relations entre mathématiques et physique. Quelle est-elle?

3. Einstein et les mathématiques

AC : D'abord, tout modèle est révisable et dépend du temps, en physique comme dans d'autres disciplines. On a bien appris la leçon, on est obligé d'être convaincu qu'un modèle de la réalité physique sera toujours tôt ou tard supplanté par un autre. C'est le côté révisable de notre perception de la nature. On peut même aller plus loin et se demander en quel sens la vérité physique dépend des questions que nous posons à la nature à travers les expériences que nous réalisons. Cependant, je voudrais insister sur le fait que dès qu'un modèle de la physique est suffisamment élaboré, la générativité des mathématiques entre en jeu : on peut alors avoir l'impression de faire de la physique en étudiant ce modèle d'un point de vue strictement mathématique. L'évolution d'Einstein est révélatrice à cet égard. Les difficultés de nature mathématique qu'il

rencontrait pour exprimer le principe de relativité générale ont modifié son attitude : il a cessé d'être le physicien pur qu'il était certainement en 1905, pour devenir mathématicien. Il a passé une très grande partie de sa vie scientifique à essayer de mettre au point une théorie qui unifie l'électromagnétisme et la gravitation. Le succès du modèle mathématique de la relativité générale était tel qu'il en était arrivé à penser que la solution de son problème se situait dans les mathématiques. En 1921, il écrivait à propos de la relativité : « Je voudrais qu'on comprenne bien que cette théorie n'est pas spéculative à l'origine; elle doit entièrement sa découverte au désir d'élaborer une théorie physique qui soit capable de rendre compte le mieux possible des faits observés. Ce n'est pas un acte révolutionnaire : l'abandon d'un certain type de concept ne doit pas être jugé comme arbitraire, mais simplement comme la conséquence directe de l'observation des faits. » Mais, en 1933, il écrit au contraire : « S'il est vrai que ce fondement axiomatique de la physique théorique ne peut provenir de l'expérience, mais doit être inventé, pouvons-nous alors espérer trouver un jour la bonne méthode? Je suis convaincu qu'il nous est possible de découvrir, par le biais de constructions mathématiques pures, les concepts et les lois les reliant les uns aux autres, lesquels doivent fournir la clé qui permettra de comprendre les phénomènes naturels... Le principe créateur se trouve dans les mathématiques ¹. »

Nous assistons en ce moment, en physique théorique, à un phénomène très semblable : à bout de ressources, un physicien théoricien en arrive à devenir mathématicien faute de mieux. Je veux parler de la théorie des cordes. Les physiciens ont cherché vers la fin des années soixante à trouver directement, sans connaître le mécanisme local des inter-

1. In Schwinger, *Einstein's Legacy*, trad. A.C., *L'héritage d'Einstein*, Pour la Science, 1988.

actions fortes, la forme mathématique d'une matrice appelée matrice S , qui détermine la probabilité pour que deux particules incidentes de moments donnés p^1 et p^2 donnent après interaction forte deux particules sortantes de moments p^3 et p^4 . Il s'agit de trouver une fonction de 4 variables p^1 , p^2 , p^3 et p^4 . L'invariance relativiste permet de la réduire à une fonction de deux variables. En posant une hypothèse simplificatrice, on parvient alors à la résoudre, et à spécifier sous forme d'intégrale une solution pour des processus invoquant même plus de 4 particules. C'est ce qu'on appelle le modèle de Veneziano. Ensuite les physiciens théoriciens ont démontré, point de départ de nombreux développements, que ce modèle décrivait en fait l'interaction non pas de particules ponctuelles, mais de petites cordes (*voir figure 11bis*).

L'intérêt de cette théorie pour les interactions fortes a cependant été de courte durée, car elle a été supplantée par la chromodynamique après la démonstration par t'Hooft de la possibilité de renormaliser les théories de jauge, la découverte de la liberté asymptotique, etc. Enfin, vers 1980, les théories des cordes ont été ressuscitées comme étant non pas des modèles des interactions fortes mais des modèles de la gravité quantique.

JPC : S'agit-il du même formalisme mathématique ?

AC : Du même formalisme mathématique. Pour te dire le changement d'échelle : en interaction forte, l'échelle standard est 10^{-13} centimètres, alors qu'en gravité c'est 10^{-33} centimètres. Il faut donc des énergies bien au-delà de toutes celles qu'on peut atteindre, de sorte qu'on sait qu'aucun phénomène expérimental qu'on puisse tester ne résultera de la théorie. Cette théorie, pour le moment, n'a de conséquences que sur un plan, je ne dirai pas purement philosophique parce que ce n'est pas vrai, mais plutôt formel. On sait que la théorie des champs peut être améliorée dans ses divergences en introduisant ces « cordes ». On remplace

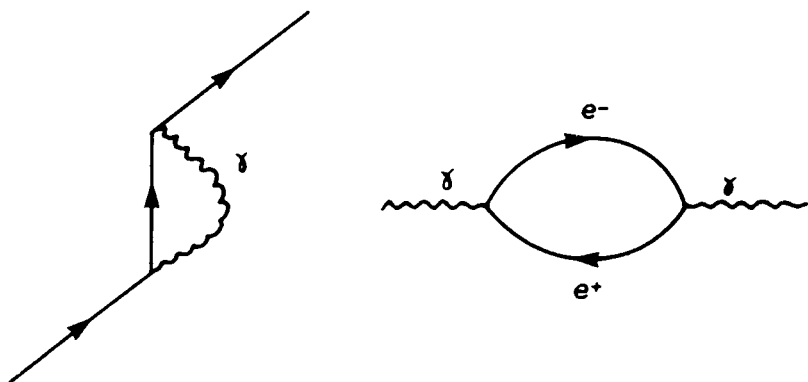


FIGURE 11

Exemples de diagrammes divergents en électrodynamique quantique.

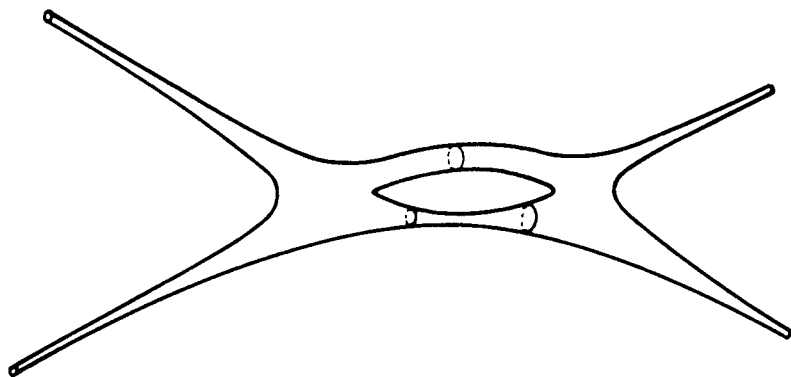


FIGURE 11bis ,

Diagramme sans singularité de la théorie des cordes.

des points par des cordes, et les particules par de toutes petites cordes qui se déplacent. On peut décrire très simplement l'intérêt que cela présente. Lorsque deux particules se rencontrent pour en recréer une, ou qu'une particule se divise en deux, on est en présence d'un processus singulier, une sorte de point dont partent trois branches (*figure 11*). Dans ce cas, il existe une singularité qui est à l'origine des divergences dont je parlais tout à l'heure et qu'on rencontre lors de l'échange d'une ou plusieurs particules virtuelles. Mais si on remplace la ligne que suivait la particule par un petit cylindre sur lequel la corde se déplace, on voit bien, en faisant des raccords de tuyauterie, que trois cylindres peuvent se rejoindre sans singularité et qu'ils sont partout parfaitement arrondis (*figure 11bis*). Que peut-on attendre de cette théorie? Parce qu'on n'a plus de singularité en ayant remplacé les trajectoires par des cylindres, c'est qu'elle soit finie au lieu d'être infinie, comme c'est le cas classiquement.

Je voudrais répéter que mon attitude personnelle à l'égard de la physique n'est pas du tout une attitude de physicien, quoique j'admire toutes les découvertes qui ont été faites, comme celle de Heisenberg, de manière très pragmatique, c'est-à-dire en partant des expériences. Les physiciens ont fait une découverte extraordinaire, la théorie des champs, mais celle-ci ne cadre pas encore de manière simple avec la partie de la réalité mathématique dévoilée jusqu'à présent. Le matériel brut existe, en quantité, et il ne s'agit plus de rechercher des résultats expérimentaux. Les mathématiques sont en retard par rapport à ce qu'elles auraient pu être pour digérer cet apport des physiciens. Il nous faut donc travailler en nous polarisant peut-être sur ces découvertes des physiciens, mais surtout travailler au sein des mathématiques pures, et non pas essayer, de manière artificielle, de faire cadrer des choses qui ne cadrent pas de manière naturelle.

JPC : L'impression que tout cela me donne c'est que le

travail des physiciens et des mathématiciens ressemble beaucoup à du « bricolage intellectuel », pour reprendre le terme cher à Claude Lévi-Strauss¹ et à François Jacob². On prend un modèle par ci, on l'applique à une observation expérimentale par là. La théorie des cordes ne colle pas pour rendre compte de la diffusion des particules. On l'abandonne! Mais la voici soudain adéquate pour améliorer la gravité quantique. Nous voilà plutôt avec du « prêt à porter » théorique qu'avec du « sur mesure ». Cela démystifie quelque peu et rend plus sympathiques ce qu'il est convenu d'appeler un peu à la légère les « sciences exactes »!

D'autre part je constate que lorsque tu parles de la relation de objets mathématiques avec les objets physiques, tu utilises le terme « cadrer » et pas « s'identifier ». En parlant de cadre, tu définis un mode de description de la réalité physique très particulier. Il me semble au contraire que si les mathématiques se trouvaient dans la nature, si la matière était organisée par les lois des mathématiques, on devrait arriver à une *identification* parfaite entre objets mathématiques et objets naturels. Or, on n'y arrive pas. Cela veut dire, selon toi, qu'elles ne s'y trouvent pas. Elles se trouvent ailleurs, mais où? Dans un état, sous une forme, que tu ne définis toujours pas. Tu arrives à une sorte de dualisme entre matière et mathématique qui rappelle le clivage que, bien entendu, je n'accepte pas, entre le corps et l'esprit.

AC : Le dualisme entre le corps et l'esprit se place sur un plan différent. Le monde physique qui nous entoure, sans être le siège de la réalité mathématique, a une cohérence certaine et difficilement explicable avec cette réalité mathématique. Comme l'a dit Einstein, si je me souviens bien, le trait le plus incompréhensible de la physique est qu'elle soit compréhensible. Il est difficile de comprendre

1. C. Lévi-Strauss, *La pensée sauvage*, Plon, 1962.

2. F. Jacob, *Le jeu des possibles*, Fayard, 1982.

que les mathématiques gouvernent l'organisation des phénomènes naturels.

JPC : Je reprends le terme « organisation des phénomènes » et j'ajoute « dans notre cerveau ».

AC : Je ne sais pas. Je ne comprends pas qu'on puisse dire « dans notre cerveau ». On pourrait aussi dire que la perception du monde extérieur se situe dans notre cerveau.

JPC : Et c'est vrai.

AC : Oui, mais nous sommes tout de même tous d'accord pour dire que le monde extérieur existe indépendamment de nous.

JPC : Oui, mais nous ne l'appréhendons que grâce à notre cerveau et à nos organes des sens.

AC : La relation avec le monde mathématique est exactement la même. Il existe en dehors de nous parce que tous les mathématiciens s'accordent sur sa structure indépendante d'une perception individuelle. D'un autre côté, il est évident qu'on peut très bien pousser quelqu'un à dire que le monde mathématique ne se réalise qu'à travers son cerveau, tout comme il ne perçoit le monde physique extérieur qu'à travers son cerveau.

JPC : Bien entendu. Je comprends. Mais je ne suis pas d'accord. En particulier lorsque tu dis « tout comme ». J'ai déjà souligné le danger qu'il y a à employer la métaphore dans de telles conditions. Analogie n'est pas démonstration. Les biologistes ont finalement une relation avec les mathématiques encore plus simple que celle des physiciens, et beaucoup moins ambiguë. La construction de modèles nécessite l'utilisation d'un appareil mathématique, et même se confond parfois avec les mathématiques, comme tu viens de le dire. Notre point de vue est peut-être moins ambitieux, mais il crée une distance plus importante. De ce fait, nous avons, je pense, une position plus juste que celle de certains physiciens.

AC : La distance est plus grande, certainement. L'im-

brication des mathématiques et de la physique explique que les physiciens ne peuvent que difficilement garder leurs distances. Oui, je suis d'accord.

4. *L'utilité des modèles mathématiques en biologie*

JPC : La croyance aux vertus explicatives du modèle mathématique se rencontre plus rarement chez les biologistes. Pour nous, les mathématiques servent principalement à deux choses. La première est l'analyse des données expérimentales...

AC : C'est de la statistique.

JPC : Oui, de l'extraction de données. Un ordinateur peut même s'en charger de manière automatique, sans faire appel au cerveau de l'expérimentateur. Les mathématiques nous aident ensuite à construire des modèles théoriques. Ceux-ci s'élaborent à partir des données de l'expérience, comme en physique. On retient des prémisses qui vont être, par exemple dans le cas de la propagation de l'influx nerveux, la mesure d'une variation de potentiel à un point donné d'un nerf et la mesure des courants transportés par les ions sodium ou potassium en fonction du potentiel. Hodgkin et Huxley ont proposé une équation ¹ qui, sur la base de ces prémisses, rend compte des bases ioniques de l'influx nerveux. Cette équation permet de décrire le phénomène, de le reconstruire à partir de données élémentaires (*voir figures 12 et 12bis*).

AC : C'est une manière de coder l'information...

JPC : Et surtout de la reconstituer.

1. A. Hodgkin et A. Huxley, *A quantitative description of membrane current and its applications to conduction and excitation in nerve*, J. Physiol., Londres, n° 117, 1952, p. 500-544.

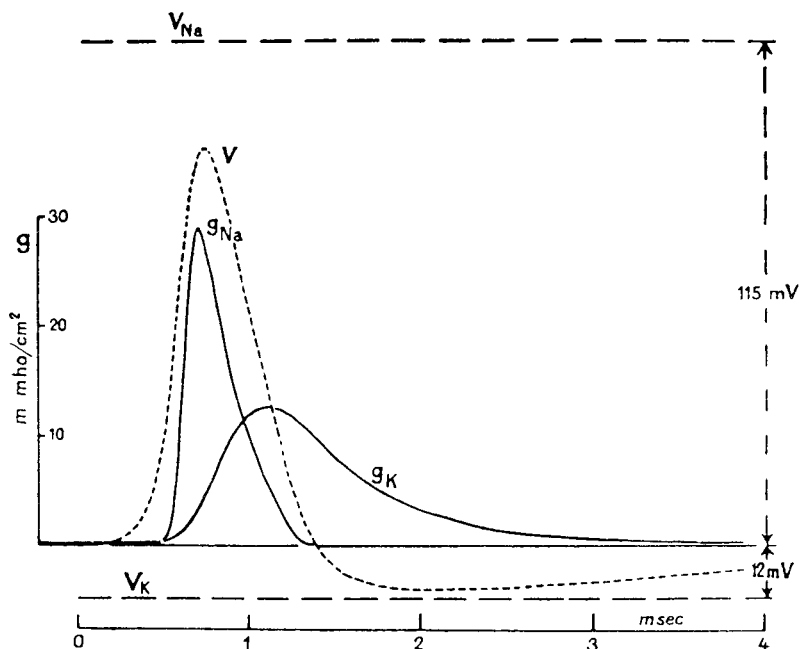


FIGURE 12

Modèle de Hodgkin et Huxley pour l'influx nerveux. L'onde de potentiel propagée est représentée par un trait discontinu (V). Celle-ci peut être décomposée en transport d'ions Na^+ vers l'intérieur de la cellule, et d'ions K^+ vers l'extérieur, exprimés ici sous forme de conductances (g^m et g^k). D'après A. Hodgkin & A. Huxley, Cold Spring Harbor, Quant. Biol, 1952, 17, 43-52.

FIGURE 12bis

Structure primaire du canal sélectif pour les ions sodium engagé dans la propagation de l'influx nerveux. Les méthodes de la génétique moléculaire ont conduit à l'identification du matériel génétique codant pour la protéine responsable du transport d'ions Na^+ , à travers la membrane du neurone, lors de la propagation de l'influx nerveux. Cette protéine se compose d'une seule chaîne de 1 820 acides aminés. La ligne inférieure est la séquence d'ADN représentée par une suite de triplets des 4 bases ATGC, la ligne supérieure est la séquence de protéine résultant de l'enchaînement d'acides aminés pris parmi les 21 acides aminés naturels, figurés ici par un code à 3 lettres. D'après Noda et al. (1984).

[illegible]

AC : Donc c'est un peu comme un langage, puisqu'un langage est fait pour reproduire...

JPC : Oui. Il permet de reproduire, mais il a aussi un caractère prédictif. Toutefois aucun biologiste, à ma connaissance, ne dira que l'équation de Hodgkin et Huxley s'identifie à l'influx nerveux, ni même qu'elle commande sa propagation. Ce n'est pas une loi mathématique de l'Univers qui dicte la propagation de l'influx nerveux, comme certains physiciens aiment à le dire lorsqu'ils parlent de leur travail!

AC : Je crois que tu as posé une question précise sur ce cas-là. Si on pousse l'analyse, disons chimique et électrique, du phénomène, je suppose qu'on parvient, grâce aux lois de la chimie, à démontrer l'équation en question.

JPC : C'est un point très important. Cette équation mathématique pourra ultérieurement être expliquée, au moins en partie, par les processus moléculaires sous-jacents. La molécule qui contient le canal sensible au voltage à travers lequel passent les ions sodium vient d'être isolée, et l'acide nucléique qui le code a été cloné et séquencé ¹. Désormais nous avons en main les mécanismes moléculaires qui déterminent la propagation de l'influx nerveux. Mais ce qui est important à retenir, c'est que l'équation mathématique n'a pas permis d'accéder directement à la structure élémentaire qui rend finalement compte du phénomène. L'accès à cette structure a été le fait d'une démarche très différente qui repose sur les méthodes de la biochimie et de la biologie moléculaire. L'équation mathématique de la propagation de l'influx nerveux se fonde sur un certain nombre de propositions portant sur les canaux postulés par le modèle. Elle

1. M. Noda, S. Shimizu, T. Tanabe, T. Takai, T. Kayano, T. Ikeda, H. Takahashi, H. Nakayama, Y. Kanaoka, N. Minamino, K. Kangawa, H. Matsuo, M. Raftery, T. Hirose, S. Inayama, H. Hayashida, T. Miyata & S. Numa, *Primary Structure of Electrophorus Electricus sodium channel deduced from cDNA sequence*, Nature 312, p. 121-27, 1984.

définit certes un certain nombre de *propriétés ioniques* élémentaires que la molécule responsable doit présenter. Mais elle ne nous dit pas du tout si ces canaux sont des protéines ou des lipides. L'équation fait appel à des phénomènes coopératifs qui ont lieu au niveau de la membrane et du transport ionique. Elle ne nous apprend pas quel va être le nombre *exact* de sous-unités engagées ou de protéines mises en jeu. Les mathématiques ont pour le biologiste un rôle prédictif certain, mais limité. Elles ne nous permettent pas d'accéder *directement* à la structure.

J'illustrerai ce point par un autre exemple, celui des lois de l'hérédité. C'est un des exemples les plus connus et les plus simples. Mendel, en étudiant la transmission héréditaire de la couleur des fleurs chez les pois, montre que celle-ci suit des lois qui se formulent par une équation mathématique d'une extrême simplicité. Les lois de Mendel ont permis d'inférer l'existence de déterminants stables et transmissibles par l'hérédité, mais n'ont certes pas permis de prévoir que c'étaient les chromosomes, et encore moins l'ADN qui étaient les supports matériels de cette hérédité.

Dans les deux cas que je viens de citer, celui de la propagation de l'influx nerveux et celui des lois de Mendel, l'équation mathématique décrit une fonction. Elle permet de cerner un comportement, mais pas d'*expliquer* le phénomène. En biologie, l'explication va de pair avec l'identification de la structure qui, sous-jacente à la fonction, la détermine. La mise en relation structure/fonction est la voie royale de la découverte, et pas seulement la description d'un processus par une équation mathématique.

AC : Je suis d'accord avec ton interprétation. En physique, de la même manière, on commence souvent par écrire des équations pour un champ moyen, dans le style de la physique du XIX^e siècle. Tant que l'on ne connaît pas la structure microscopique sous-jacente, on ne peut guère démontrer ces équations. Mais quand la théorie est suffi-

samment élaborée, la générativité des mathématiques entre en jeu. Mon exemple favori est emprunté à Heisenberg. Guidé par les résultats de la spectroscopie expérimentale comme le principe de combinaison de Ritz Rydberg, Heisenberg a compris que l'algèbre des quantités observables pour un système atomique devait être une algèbre non commutative, l'algèbre des matrices. Cette seule observation et un peu de mathématiques engendrent l'équation de Schrödinger qui explique les nombres mystérieux (différences entre les inverses des carrés de deux nombres entiers) qui gouvernent les régularités des raies spectrales de l'atome d'hydrogène. Avec le principe d'exclusion de Pauli et des mathématiques plus élaborées, on arrive ensuite à analyser l'équation de Schrödinger pour un atome à n électrons.

JPC : Et à décrire complètement le tableau de Mendéléev.

AC : Ce qui est extraordinaire. Dans la modélisation d'un phénomène, on peut voir deux étapes. Celle des physiciens du XIX^e siècle d'abord, quand ils regardaient l'écoulement des flux ou lorsqu'ils décrivaient des phénomènes de manière macroscopique. Ensuite, lorsqu'on a mieux compris la structure microscopique de la matière, on est parvenu à utiliser la générativité des mathématiques, qui permet de trouver qu'en fait il y a peu de possibilités et de prévoir même la chimie (*voir figure 13*).

JPC : Mais cet aspect génératif ne tient-il pas précisément, comme tu viens de le dire, au fait qu'on accède à un niveau inférieur et que des régularités apparaissent qui, ensuite, sont appliquées de manière universelle ?

AC : Bien sûr. Tant qu'on n'a pas réussi à accéder à ce niveau qui va au-delà du champ moyen, je pense que l'aspect génératif est limité.

JPC : C'est presque la même chose que l'exemple de la biologie. L'équation de Hodgkin et Huxley est généralisable. Elle a un aspect prédictif. Mais, dès qu'on accède à l'analyse

Z	Élément	$W_i(\text{ev})$	Configuration électronique									
			1s	2s	2p	3s	3p	3d	4s	4p	4d	4f
1	H <i>hydrogène</i>	13.6	1									
2	He <i>hélium</i>	24.6	2									
3	Li <i>lithium</i>	5.4	remplis (2)	1								
4	Be <i>beryllium</i>	9.3		2								
5	B <i>bore</i>	8.3		2	1							
6	C <i>carbone</i>	11.3		2	2							
7	N <i>azote</i>	14.5		2	3							
8	O <i>oxygène</i>	13.6		2	4							
9	F <i>fluor</i>	17.4		2	5							
10	Ne <i>néon</i>	21.6		2	6							
Nombre d'électrons de chaque état												
11	Na <i>sodium</i>	5.1				1						
12	Mg <i>magnésium</i>	7.6				2						
13	Al <i>aluminium</i>	6.0				2	1					
14	Si <i>silice</i>	8.1	— remplis —			2	2					
15	P <i>phosphore</i>	10.5				2	3					
16	S <i>soufre</i>	10.4	(2)	(8)		2	4					
17	Cl <i>chlore</i>	13.0				2	5					
18	A <i>argon</i>	15.8				2	6					
19	K <i>potassium</i>	4.3									1	
20	Ca <i>calcium</i>	6.1									2	
21	Sc <i>scandium</i>	6.5								1	2	
22	Ti <i>titanium</i>	6.8								2	2	
23	V <i>vanadium</i>	6.7								3	2	
24	Cr <i>chrome</i>	6.8								5	1	
25	Mn <i>manganèse</i>	7.4	(2)	(8)		(8)				5	2	
26	Fe <i>fer</i>	7.9								6	2	
27	Co <i>cobalt</i>	7.9								7	2	
28	Ni <i>nickel</i>	7.6								8	2	
29	Cu <i>cuivre</i>	7.7								10	1	
30	Zn <i>zinc</i>	9.4								10	2	
31	Ga <i>gallium</i>	6.0									2	1
32	Ge <i>germanium</i>	7.9									2	2
33	As <i>arsenic</i>	9.8									2	3
34	Se <i>sélénium</i>	9.7	(2)	(8)		(18)					2	4
35	Br <i>brome</i>	11.8									2	5
36	Kr <i>krypton</i>	14.0									2	6

FIGURE 13
Début du tableau périodique des éléments.

des canaux ioniques individuels et des molécules dont l'activité collective constitue l'influx nerveux, un nouvel ensemble de règles et de prédictions apparaît. Celles-ci se formulent sous une forme mathématique nouvelle qui s'applique à d'autres systèmes, aux canaux sélectifs pour le calcium ou à ceux sensibles à des neurotransmetteurs.

AC : Oui, je suis tout à fait d'accord. Je voudrais tout de même faire une critique générale en ce qui concerne le type de mathématiques utilisé dans ce genre de modélisation. Les mathématiques qu'il implique tournent toujours autour d'équations aux dérivées partielles, ou, au mieux, de modèles de mécanique statistique. Dans les deux cas, et comme dans la plupart des modèles de la physique, le principe directeur est la notion fondamentale de localité des interactions. Même des interactions qui ont une apparence de non-localité, comme l'attraction newtonienne, redeviennent locales lorsqu'on introduit les champs convenables. Ce principe de localité des interactions est une règle d'or de la physique moderne dont l'outil essentiel est la manipulation des Lagrangiens. Pourtant, il ne me paraît pas évident, au moins *a priori*, que les seules mathématiques intéressantes et utiles pour un biologiste passionné par le fonctionnement du cerveau soient celles dont je viens de parler. Il serait bon qu'au moins les rudiments de notions comme celle de topologie combinatoire soient sinon utilisés, du moins connus des biologistes.

JPC : Cela arrivera dans la suite de cet exposé.

AC : C'est exactement la raison pour laquelle j'étais très intéressé par notre rencontre. En biologie, on utilise les mathématiques comme un langage. Si, par exemple, vous tenez une courbe de réponses, il est bien évident qu'il est plus facile de la transmettre si l'on a une fonction mathématique simple qui permet de la décrire que si l'on doit retracer la courbe et en isoler les paramètres. C'est simplement une manifestation de la jeunesse de la biologie. Si l'on

regarde le développement de la physique, on s'aperçoit qu'au début on arrivait à formaliser certains phénomènes, à les transmettre sous forme de fonctions mathématiques. C'était par exemple le cas pour la découverte de Planck. Mais au bout d'un moment, les mathématiques, parce qu'elles sont génératives, permettent d'ajouter du nouveau. Et pas seulement parce que les équations permettent la prédiction. La cohérence interne aux mathématiques se manifeste, comme dans le cas de l'atome de l'hydrogène, et permet, en vertu de critères de simplicité, d'esthétique mathématique, d'avoir une très bonne intuition de ce qui devrait être vrai dans des cas où l'on n'a pratiquement pas fait d'expériences à l'avance, et ensuite de vérifier si, oui ou non, ça marche. Je suis donc très optimiste quant au rôle génératif que les mathématiques pourront éventuellement jouer en biologie. Je pense qu'ultérieurement, peut-être pas encore maintenant, quand on aura réussi à voir quelle partie des mathématiques peut être le mieux utilisée et peut le mieux cadrer avec la biologie, leur générativité sera extrêmement utile.

5. Auscultation de la mécanique quantique

JPC : Je voudrais revenir à la mécanique quantique et aux conclusions qu'on peut tirer de l'usage grossier, fruste, mais après tout assez pertinent, des mathématiques par les biologistes. Nous nous transformons en mathématiciens de temps en temps, ou bien nous coopérons avec des mathématiciens, pour essayer de trouver, en quelque sorte, l'habillage mathématique qui convient le mieux aux phénomènes biologiques qui nous intéressent. Il n'est donc absolument pas question pour nous d'identifier le réel biologique à des objets mathématiques. Nous cherchons seu-

lement à construire des objets mathématiques qui s'ajustent aux objets naturels. On raisonne, on résout, on élabore modèle après modèle, on se réfère à une littérature qui témoigne déjà de multiples essais et erreurs. Mais finalement, que fait-on? On *sélectionne* le modèle qui s'adapte le mieux. Nous avons donc une vision extrêmement pragmatique et concrète des mathématiques. Nous ne retenons d'elles que ce qui est le plus adéquat à la réalité naturelle. Les mathématiques sont pour nous des objets de pensée. Ni plus, ni moins.

Cela m'incite à évoquer à nouveau avec toi la mécanique quantique, un domaine de la physique que je connais mal. J'ai le sentiment que les physiciens travaillent dans un domaine où ils éprouvent des difficultés à se représenter ce qui se passe à une échelle très différente de celle où notre cerveau et nos organes des sens fonctionnent (*voir figure 14*).

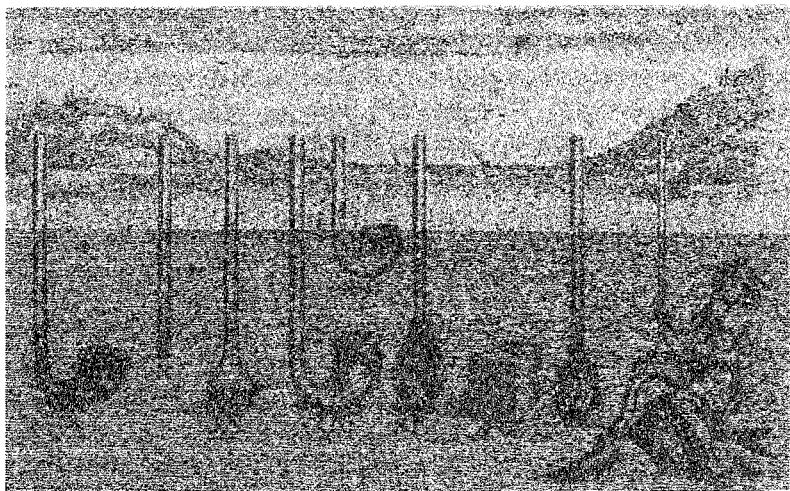


FIGURE 14

Gravure extraite du Traité de l'équilibre des liqueurs de Blaise Pascal (1664, 2^e édition, C. Desprez, Paris). Elle pose le problème de la mesure en physique et de sa relation avec l'expérimentateur. Celui-ci est curieusement représenté en dessous du niveau de l'eau!

Et quand un physicien nous dit que les lois de la physique quantique imposent une indétermination fondamentale, j'emploie à dessein ces termes que l'on retrouve sous leur plume, on peut se demander s'ils ne font pas en fait une erreur épistémologique grave...

AC : Tu veux dire « de l'ordre du langage ».

JPC : Qui est d'identifier la nature au modèle qu'ils construisent pour la décrire. On peut se demander s'ils prennent sérieusement en considération, non seulement l'instrument de mesure et le regard de l'observateur, mais aussi leur propre fonctionnement cérébral et l'aptitude de celui-ci à capturer les phénomènes à une échelle où l'expérience ordinaire et le sens commun ne s'appliquent plus. Qu'en penses-tu ?

AC : J'ai été moi-même confronté à cette question d'une indétermination fondamentale. C'est pourquoi je peux te répondre. Se pose d'abord, mais ce n'est pas l'essentiel, un problème de langage. Lorsqu'on parle d'une particule et qu'on l'imagine comme un point matériel qui a une position et une vitesse déterminées, on sait que ces termes sont mauvais. Si par exemple on veut se donner une image mentale de l'électron qui tourne autour du noyau de l'atome d'hydrogène, il vaut beaucoup mieux penser à la fonction d'onde déterminée par l'équation de Schrödinger et le niveau d'énergie que de penser à un système planétaire. De plus, pour un atome plus complexe, comme l'atome d'hélium, qui a deux électrons autour de son noyau, l'image mentale est beaucoup plus difficile, car la fonction d'onde, dont on pouvait, dans le cas d'un seul électron (*voir figure 15*), au moins visualiser la taille dans l'espace, est maintenant une fonction de *deux* variables dans l'espace, c'est-à-dire une fonction sur un espace à six dimensions. Le pas suivant consiste à comprendre que bien que le langage des particules soit inadapté, il permet de formuler des questions, et que la nature donne une réponse à ces questions. Pour prendre

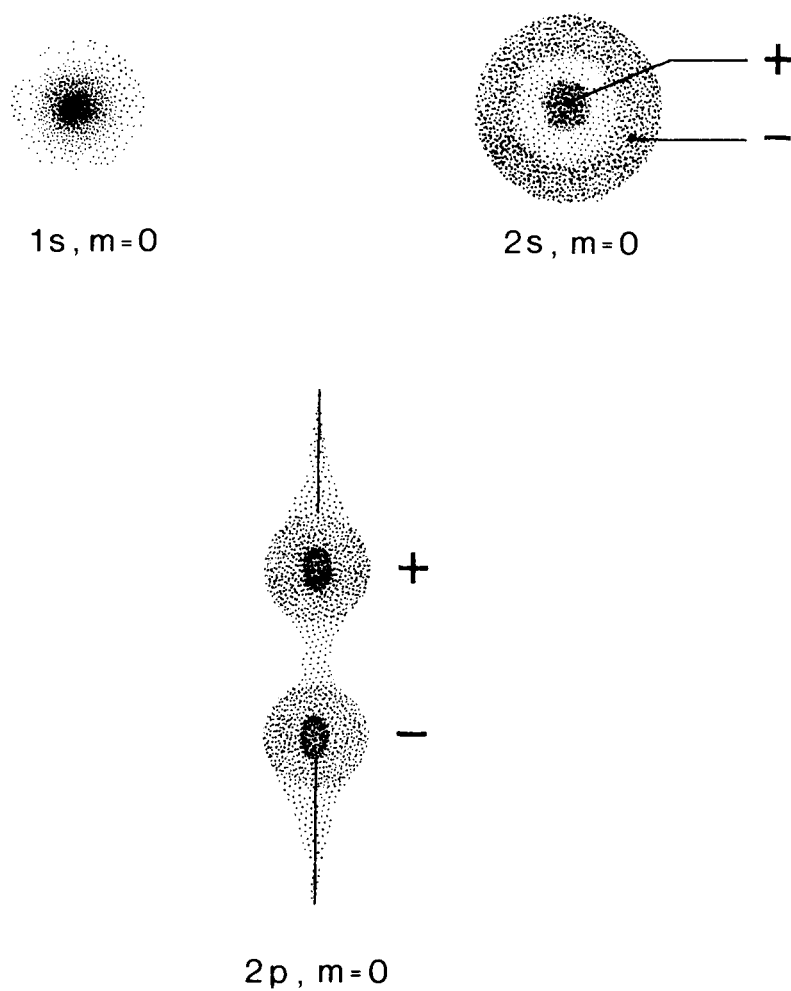


FIGURE 15

Exemples de fonction d'onde de l'électron dans l'atome d'hydrogène.

un exemple concret, on peut considérer une source discrète de particules, d'électrons par exemple, qui émet de temps en temps un électron dans la direction d'une fente très fine, créant ainsi un phénomène de diffraction. On peut décrire le système à partir de fonctions d'onde, et prévoir une figure de diffraction sur un écran placé après la fente, le long du parcours de l'électron. Si le langage des particules était entièrement mauvais, on en aurait alors la preuve : l'électron devrait se transformer en un nuage à cause du phénomène de diffraction. Or, ce n'est pas le cas. Si l'on fait l'expérience, on obtient chaque fois un impact à un endroit précis de l'écran. L'électron reste donc une particule. C'est précisément au cours d'expériences de ce type que l'indéterminisme fondamental dont tu parlais apparaît. En effet, chaque fois qu'un électron est émis par la source, on mesure un impact à un endroit x de l'écran d'arrivée (*voir figure 16*). Mais, et c'est là le point crucial, le résultat expérimental « la source émet un électron qui arrive au point x de l'écran » n'est pas reproductible. Cela n'a rien à voir avec la précision avec laquelle on donne x . Même le résultat expérimental « la source émet un électron qui arrive dans la moitié supérieure de l'écran » n'est pas reproductible. On ne pourra jamais rendre les données initiales de l'expérience suffisamment précises pour que le résultat final soit à coup sûr le même. A la deuxième expérience, on aura même seulement une chance sur deux que le résultat final soit le même. Quelle que soit la précision de l'appareil qui envoie les électrons, on ne peut pas répéter l'expérience avec le même résultat. Seule est reproductible la fréquence, la probabilité, avec laquelle l'électron va arriver à la sortie. La seule donnée reproductible, c'est une certaine densité, une certaine courbe de fréquence d'impact à l'arrivée, sur l'écran d'arrivée. Elle aura la forme d'une courbe de diffraction et permettra de dire que l'électron est arrivé avec telle probabilité à tel ou tel endroit.

JPC : L'existence d'un indéterminisme fondamental n'est pas prouvée pour autant. Tu pourrais avoir la même chose dans une situation plus macroscopique, du fait par exemple du mouvement brownien...

AC : Si. Il faut bien comprendre qu'on fait erreur si l'on considère comme un résultat reproductible le fait que l'électron a fait « tic » à tel endroit de l'écran. Aucune théorie ne peut prévoir le résultat de cette expérience, parce qu'il n'est pas reproductible. Lorsqu'on veut faire de la physique, on doit définir ce qu'est un phénomène physique.

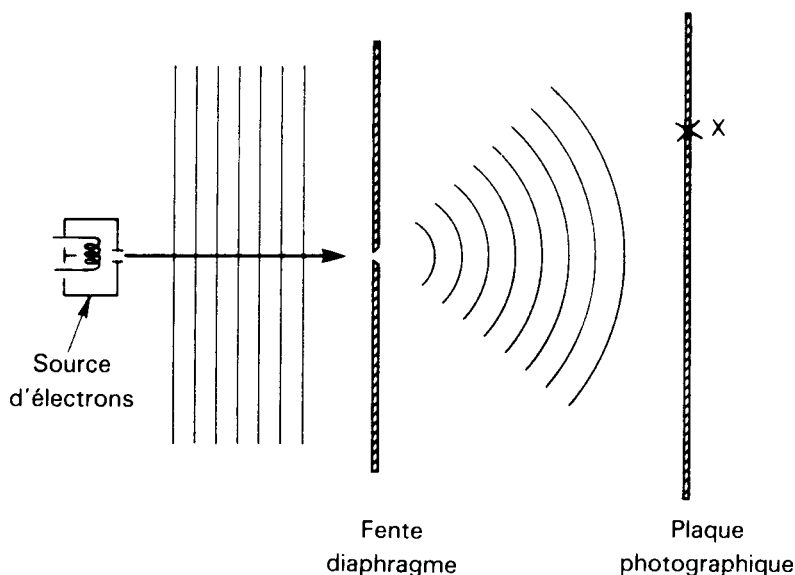


FIGURE 16
Phénomène de diffraction.

A partir du moment où l'on s'en donne une définition cohérente, il n'y a absolument rien de déroutant, aucun paradoxe dans la mécanique quantique. La théorie cadre parfaitement avec son modèle mathématique. Quelle définition peut-on donner d'un phénomène physique? On ne peut parler de phénomène physique qu'à propos de résultat expérimental reproductible. Un phénomène physique est donc le résultat d'une expérience dans laquelle, si l'on spécifie les données initiales, ne serait-ce qu'aux expérimentateurs d'un autre laboratoire, le résultat sera identique. Si au contraire on ne peut pas leur transmettre de manière précise les données initiales, de telle sorte que le résultat expérimental soit le même, le phénomène que l'on est en train d'étudier n'est pas un phénomène physique. De ce fait, il ne peut être prévu par une théorie.

JPC : Donc il n'y a pas d'indétermination fondamentale. Le fait qu'un électron ait à un moment tel point d'impact, et à un autre tel autre, pourrait finalement s'expliquer, un jour ou l'autre, sous une forme déterministe.

AC : Non. On sait que ce phénomène n'est pas explicable par ce qu'on appelle les « variables cachées ».

JPC : Parce que l'hypothèse des variables cachées est issue d'un modèle particulier. Mais peut-être existe-t-il un autre modèle auquel les physiciens n'ont pas encore pensé.

AC : Non. L'hypothèse des variables cachées n'est pas compatible avec le modèle actuel de la mécanique quantique, modèle dont la seule justification est son incroyable succès. On peut imaginer une suite d'expériences testant les inégalités de Bell qui montrent que l'interprétation des variables cachées est incompatible avec le modèle actuel. Mon point de vue est très simple. Il y a certains résultats qui sont des résultats de physique parce qu'ils sont reproductibles. Et il y en a d'autres qui n'en sont pas, parce qu'ils ne sont pas reproductibles.

JPC : Je m'interroge sur l'emploi du terme « non-repro-

ductibilité ». Si on enregistre un neurone lorsqu'on stimule l'œil, on obtient une rafale d'impulsions. Mais, d'une expérience à l'autre, on n'enregistrera pas nécessairement le même nombre d'impulsions...

AC : Bien entendu. Par contre, on pourra en extraire une loi qui, elle, sera reproduite, et c'est la seule qui compte.

JPC : Exactement. Après un certain nombre d'expériences se dégagera une réponse moyenne beaucoup plus reproductible. Et bien des biologistes diront qu'il existe un « déterminisme » de la genèse de cette réponse, du récepteur sensoriel au neurone, en dépit des « fluctuations » qui interviennent dans la réponse. J'emploie le terme « fluctuations », parce qu'il s'agit d'une variabilité due aux modalités plus fines de la transmission du signal au niveau des synapses, de la géométrie des contacts entre cellules nerveuses, etc., que l'expérimentateur ne contrôle pas dans une expérience de ce type. Il existe de fait plusieurs niveaux de variabilité dans le transfert d'informations à l'intérieur du système nerveux. Ils n'ont rien de mystérieux¹. Si on applique ce mode de pensée aux processus qu'étudient les physiciens, on peut se demander si ceux-ci n'ont tout simplement pas encore trouvé le modèle qui apporte une explication plus profonde. Le modèle des variables cachées ne marche pas. Je reste cependant hésitant devant un résultat négatif ou une interprétation qui échoue. Peut-être pourra-t-on, un jour, aboutir à une explication plus rationnelle.

AC : Dans le cas de l'œil, on sait que, si l'on répétait l'expérience avec les mêmes données initiales, on obtiendrait la même rafale d'impulsions.

JPC : Oui. C'est théoriquement possible, mais pratiquement impossible.

1. Y. Burnod & H. Korn, *Consequences of Stochastic Release of Neurotransmitters for Networks Computations in the Central Nervous System*, Proc. Nat. Acad. Sci. USA, 86, 352-356, 1959.

AC : Alors qu'en mécanique quantique, c'est impossible à cause de la théorie, et c'est cela la différence.

JPC : Je suis mal ton raisonnement. Cela paraît théoriquement impossible. Mais n'est-ce pas parce que la théorie adéquate n'a toujours pas été découverte ?

AC : Non. Malheureusement, c'est un point très délicat. Mon point de vue est net. Certains résultats d'expériences pourraient être considérés comme des phénomènes physiques. Or ils n'en sont pas, car ils ne sont pas reproductibles. Aucune théorie ne peut espérer prévoir un phénomène si celui-ci n'est pas reproductible.

JPC : Mais il n'est pas exclu qu'il ne le devienne. On a l'impression que les physiciens sont obnubilés par une théorie qui rend trop bien compte des phénomènes naturels pour qu'ils cherchent à en trouver une autre. Sauf s'ils sont préoccupés par le problème posé et prennent le risque d'aller plus au fond des choses.

AC : La théorie, telle qu'elle est, montre que cet indéterminisme fondamental se manifeste, dès qu'on fait deux mesures successives de quantités observables qui ne commutent pas, et le principe d'incertitude de Heisenberg précise quantitativement cet indéterminisme. Ce principe a donc une validité théorique, mais aussi expérimentale. La mécanique quantique pose ainsi au niveau expérimental le problème d'expériences, de résultats expérimentaux non reproductibles. Si, dans l'expérience de l'électron, je ferme la fente et mesure le recul de l'écran dû à l'impact de l'électron, ce phénomène est parfaitement reproductible et parfaitement explicable par la physique. C'est la règle de la conservation du moment. Si je dis : « Lorsque j'ai lancé un nombre considérable d'électrons, la probabilité d'arriver à tel endroit de la plaque a telle valeur », c'est un phénomène parfaitement reproductible et explicable par la théorie. Par contre, si je dis : « L'électron est arrivé à tel endroit de

l'écran », ce n'est pas un résultat expérimental reproductible.

JPC : Tu entends par phénomène physique un phénomène reproductible. Cela veut dire qu'il faut définir les conditions expérimentales de telle sorte qu'il devienne reproductible. Si on connaissait les événements qui déterminent le déplacement de l'électron vers le haut ou vers le bas, on rendrait ce phénomène reproductible.

AC : On ne peut pas faire ce que tu dis sans changer les résultats, même au niveau de la statistique.

JPC : Je dirais alors que la théorie est mauvaise.

AC : Oui, mais elle explique précisément les phénomènes expérimentaux reproductibles. Et comme je le disais plus haut, sous une forme plus élaborée, elle prévoit, par exemple, ce qu'on appelle le moment magnétique anormal d'un électron, avec la précision de l'épaisseur d'un cheveu sur la distance Paris-New York.

JPC : Il reste un niveau sous-jacent inexpliqué, auquel les théoriciens n'ont pas encore eu « mentalement » accès. Le modèle d'Hodgkin et Huxley est parfaitement apte à rendre compte des phénomènes électriques de l'influx nerveux en termes de transport ionique. Il n'a néanmoins pas directement conduit à identifier les canaux ioniques sous-jacents, ce qu'a réussi le biologiste moléculaire qui use de méthodes radicalement différentes de celles employées par Hodgkin et Huxley dans leurs expériences.

AC : Je vais quand même te donner une image pour te montrer comment on peut se sortir du malaise dû à l'indéterminisme apparent de la mécanique quantique. Si on limite le champ d'investigation du physicien aux phénomènes reproductibles, on aboutit à un tout parfaitement cohérent, mais l'impossibilité de prévoir à quel endroit l'impact de l'électron aura lieu est une forme de renoncement très frustrant pour le théoricien. Prenons une image bien connue des physiciens : celle des univers parallèles de

H. Everett ¹. Tout se passe comme si tous les événements possibles pouvaient se produire, comme si l'électron pouvait aller à n'importe quel endroit de la plaque. Mais chacun de ces choix signifierait une bifurcation d'un univers à un univers parallèle. Supposons par exemple, pour simplifier, que l'on fasse une mesure avec deux résultats possibles. Son résultat crée alors une bifurcation entre deux univers parallèles. On se trouvera sur l'un ou l'autre selon que l'une ou l'autre des possibilités se sera réalisée. La cohérence de ces univers parallèles est suffisante pour qu'en statistique le résultat moyen soit le même. Chaque résultat d'expérience dépend de l'univers parallèle particulier vers lequel nous bifurquons. En lui-même, il n'est pas reproductible.

JPC : L'idée paraît intéressante. Mais la confusion savamment entretenue entre irreproductibilité et indétermination suggère que la plupart des théoriciens se trouvent face à un échec que leur inconscient refuse d'avouer.

AC : On ne peut espérer prédire théoriquement un résultat non reproductible. La qualité primordiale d'une expérience physique, je pense que tout le monde est d'accord sur ce point, est sa reproductibilité. Si elle n'est pas reproductible, elle n'a pas de contenu physique. L'échec n'est pas théorique, mais expérimental. Lorsqu'on règle l'appareil expérimental, on n'arrive pas à savoir comment préciser les données initiales pour connaître à l'avance le point d'impact de l'électron. Le principe d'incertitude de Heisenberg démontre qu'il est à jamais impossible d'y parvenir, et que le même indéterminisme se manifeste dès qu'on mesure successivement deux quantités observables qui ne commutent pas, comme dans une expérience de Stern Gerlach.

JPC : Tu pourrais peut-être contrôler un paramètre physique qu'on ignore. C'est donc à la fois théorique et expé-

1. H. Everett, « *Relative State* » *Formulation of Quantum Mechanics*, *Reviews of Modern Physics*, 29, 454-62 (1957).

rimental. Il serait piquant que le mathématicien suggère une expérience aux physiciens!

AC : Il est quand même très important de voir que cet échec est un échec expérimental. C'est ce qui explique que j'aime prendre cet exemple. Simplement parce qu'on croit que seule la théorie est en jeu, alors que l'expérience, elle aussi, est concernée : on est en présence d'un résultat expérimental qu'on ne sait pas reproduire.

JPC : Une bonne expérience est plus difficile à faire qu'une théorie médiocre. Et cette discussion fait apparaître que l'indétermination mystérieuse dont parlent un certain nombre de physiciens n'a pas tellement de sens. Il faudrait plutôt accepter l'idée que l'état de nos connaissances ne nous permet pas encore de maîtriser ces notions, à la fois sur le plan expérimental et sur le plan théorique. Il me paraît difficile d'accepter l'ignorance comme une loi de la nature...

AC : Effectivement, dans le corps de doctrine actuel, parfaitement cohérent, on parvient à rendre compte des résultats expérimentaux qui sont reproductibles. Je ne vois guère comment on pourrait rendre compte de résultats qui ne le sont pas. Il est très difficile d'admettre qu'au niveau microscopique, au niveau du quantique, il y ait des phénomènes qui ne soient pas reproductibles. Il n'empêche que c'est un fait. Son impact philosophique est difficile à apprécier. Ce qui est incompréhensible ici, c'est que la nature, au niveau atomique, soit imprédictible. Même la « réalité » physico-chimique est plus subtile qu'il n'y paraît.

IV

Le mathématicien neuronal

1. *L'illumination*

JEAN-PIERRE CHANGEUX : De nos jours, peu de mathématiciens se préoccupent du cerveau. J'ai remarqué dans le livre de Dieudonné *Pour l'honneur de l'esprit humain* – dont le titre rappelle un peu le *Ad maiorem dei gloriam* d'Ignace de Loyola – que le mot « cerveau » n'apparaît que très rarement. En tout cas pas dans un but explicatif. Bien au contraire. Il écrit par exemple : « L'activité rationnelle d'un cerveau créateur n'a jamais eu d'explication rationnelle, en mathématiques pas plus qu'ailleurs ¹. » Dans ce texte, que j'aime et que j'ai lu avec beaucoup d'intérêt, Dieudonné examine l'évolution des mathématiques d'une manière tout à fait indépendante du cerveau, un peu comme les historiens de l'art s'intéressent à l'évolution de la peinture et de la sculpture sans vouloir se rendre compte que nous voyons

1. P. 38.

plus avec notre cerveau qu'avec notre œil! Il est bon de rappeler que le mathématicien fait tout de même des mathématiques avec son cerveau et qu'il ne peut pas en être autrement!

ALAIN CONNES : Je partage entièrement cette idée. Le cerveau est un outil matériel et il est fondamental de comprendre son fonctionnement dans le travail du mathématicien.

JPC : On trouve toutefois chez quelques mathématiciens du passé, comme Poincaré ou Hadamard, des préoccupations proches des nôtres. Hadamard, dans son excellent *Essai sur la psychologie de l'invention dans le domaine mathématique*¹, s'interroge sur l'inconscient et sur ses couches successives dans la création mathématique. Il cite *De l'Intelligence* de Taine, ce philosophe qui daignait encore prendre en compte les données de la science, en particulier des sciences du système nerveux, préoccupation qui, depuis Sartre, Foucault et leurs successeurs, a été oubliée de beaucoup de nos philosophes contemporains, au profit de la psychanalyse, bien que de notables exceptions commencent à faire entendre leur voix².

Hadamard décrit son travail de mathématicien d'une manière que je trouve très intéressante. Il distingue tout d'abord un travail préparatoire, qui comporte – j'insiste – des échecs et des erreurs, celles-là mêmes que le mathématicien oublie pieusement de mentionner lorsqu'il présente ses résultats sous une forme en général « bien digérée ». Se rapprochant des analyses de Poincaré, il s'efforce de cerner ces tentatives pour « gouverner l'inconscient », et distingue dans la « création mathématique » plusieurs étapes, qu'il appelle « préparation », « incubation », et « illumination ». Il

1. J. Hadamard, *Essai sur la psychologie de l'invention dans le domaine mathématique*, Gauthier-Villars, Paris, 1952.

2. Cf. *L'Age de la science*, n° 2, *Epistémologie*, sous la direction de P. Jacob, Odile Jacob, Paris, 1989.

souligne également l'emploi des signes ainsi que de l'imagerie mentale, et fait alors référence à un psychologue de l'époque, Binet, qui, comme Taine, s'est beaucoup intéressé aux expériences d'imagerie, à la suite des associationnistes anglais. Il est remarquable que cet intérêt pour l'imagerie mentale ait récemment refait surface en psychologie expérimentale, avec des auteurs comme Kosslyn, Shepard, et Denis¹ en France. On retrouve là une préoccupation commune aux psychologues et aux neurobiologistes. Car l'image mentale ne doit pas être prise dans un sens évanescent ou immatériel, mais au contraire, comme une activité cérébrale concrète et bien définie. Hadamard signale que, lors du travail préparatoire, quand les images commencent à surgir dans le cerveau du mathématicien, parfois, une soudaine *illumination* envahit son cerveau et sa sensibilité. Elle constitue une étape importante dans le travail de la création mathématique. Mais une troisième étape suit nécessairement. Plus consciente que la précédente, elle consiste en vérifications et en définitions qui permettent d'exposer avec précision un raisonnement, un théorème ou une démonstration. Cette dernière étape fait intervenir raisonnement et jugement.

La démarche d'Hadamard est de type introspectif. De ce fait, elle est souvent critiquée par les psychologues, les philosophes et bien sûr les neurobiologistes, parce qu'elle est subjective. Toutefois elle est intéressante parce qu'elle donne lieu à un récit qui possède une forme d'objectivité puisqu'il est reproductible d'un mathématicien à l'autre. Que penses-tu de cette description, commune à Hadamard et à Poincaré, de la création mathématique?

AC : J'ai moi-même vécu – du moins, je le crois – des expériences de ce type. La première phase, l'incubation, consiste en une approche fondée sur des connaissances déjà

1. M. Denis, *Image et cognition*, PUF, Paris, 1989.

acquises : progressivement, on en vient à se concentrer sur un objet de pensée bien précis. On essaie de focaliser sa pensée en préparant le terrain, en s'entourant de choses connues. La troisième, la vérification, commence une fois que l'illumination a eu lieu. Le processus de vérification est très douloureux, parce qu'on a peur de s'être trompé. C'est en fait la phase la plus angoissante, parce qu'on ne sait jamais si son intuition est juste... c'est un peu comme dans les rêves, l'intuition se trompe très facilement. Je me souviens d'avoir passé un mois à vérifier un résultat : je reprenais la démonstration dans ses moindres détails, jusqu'à l'obsession, tâche qui pourrait à la rigueur être confiée à un calculateur électronique qui vérifierait la logique du raisonnement. Au contraire, au moment où elle a lieu, l'illumination implique une part considérable d'affectivité, de sorte qu'on ne peut rester passif ou indifférent. La rare fois où cela m'est réellement arrivé, je ne pouvais m'empêcher d'avoir les larmes aux yeux. J'ai souvent observé la chose suivante : une fois la première étape de préparation franchie, on se heurte à un mur. L'erreur à ne pas commettre consiste à attaquer cette difficulté de manière frontale. Il faut procéder indirectement, procéder à côté. Si l'on pense trop directement au problème, on épuise assez vite les outils accumulés au cours de la première phase et on se décourage. Il faut libérer la pensée, de telle sorte que le travail subconscient puisse se produire. Par exemple, lorsqu'on fait des calculs algébriques relativement élémentaires, mais longs, cette durée, au cours de laquelle la pensée directe est relativement peu focalisée, est très propice à l'intervention du subconscient. Le mathématicien doit évidemment disposer d'une sérénité suffisante. On peut parvenir ainsi à une sorte d'état contemplatif qui n'a rien à voir avec la concentration d'un étudiant en mathématiques qui passe un examen. Au mieux, un étudiant qui utiliserait cette technique sortirait d'un examen en disant : « J'ai raté mon examen, mais j'ai

trouvé une idée sur laquelle j'ai envie de travailler pendant longtemps. » Ce qui est frappant, c'est l'importance, quand je parle de procéder indirectement, de l'éloignement apparent entre le problème initial et le champ d'investigation du moment.

JPC : Bien entendu. Pendant toute cette période, ton cerveau est en pleine évolution. Tu fais des hypothèses, tu crées des ébauches...

AC : Mais pas de ce problème-là.

JPC : Si tu contournes le problème, comment la solution peut-elle apparaître si soudainement ?

AC : C'est assez difficile à décrire. L'expérience montre que si l'on s'attaque à un problème directement, on épuise très vite toutes les ressources de la « pensée directe », rationnelle. On cerne la difficulté, mais si l'on ne parvient pas à se libérer, en général on ne résout pas le problème, contrairement à ce qui se produirait s'il s'agissait d'un problème d'examen, s'il n'y avait que des opérations automatiques à effectuer. Ce stade correspond à peu près à la connaissance que les mathématiciens ont d'un problème donné. Ils cernent facilement le problème, ils définissent exactement la difficulté, mais au-delà, la pensée directe n'aide plus. On ne progresse que si l'on dispose d'une stratégie, même implicite, qui consiste à réfléchir à des questions annexes *a priori* sans relations avec le problème lui-même.

JPC : Sont-elles très différentes ou voisines ?

AC : Elles peuvent être très différentes.

JPC : S'agit-il simplement d'occuper la mémoire de travail et de laisser se développer plus en profondeur un travail inconscient qui suppose une contribution plus importante de la mémoire à long terme ? Ou est-ce, au contraire, une sorte de procédure d'association qui prend du temps parce que les éléments qui vont se rejoindre font partie de contextes très différents ? J'ai cru comprendre que circuler autour du problème permettait d'évoquer des objets mathématiques

sans relation directe avec lui. Par combinaison, ils conduisent à la solution, ou bien évoquent, par un biais quelconque, en puisant dans la mémoire à long terme, une représentation plus adéquate à la question posée. S'agit-il d'un procédé d'occultation de la pensée rationnelle, d'atténuation de la conscience, qui laisse des représentations internes « incongrues » se manifester et permet à des objets mathématiques de s'associer « contre nature » ? Retrouves-tu incorporés, sous forme de « chimère », à la solution finale ces éléments d'une réflexion parallèle ? Je t'expliquerai bientôt les raisons de ma question.

AC : Je ne parle que de ma propre expérience. Bien que participant du même domaine, ma réflexion était *a priori* disjointe. Elle aboutissait à la résolution du problème lui-même sans qu'à aucun moment elle n'ait été en apparence guidée par celui-ci.

JPC : Elle était quand même au fond de toi.

AC : Probablement, mais je n'en étais absolument pas conscient. Je me posais une autre question, qui évoluait et m'amenait à la résolution du premier problème.

JPC : Je repose ma question : lorsque tu utilises, pour résoudre un problème, une connaissance connexe ou distincte, ce matériau mathématique persiste-t-il sous une forme « recombinate » dans la solution finale ?

AC : C'est difficile à dire. Le problème était le suivant : démontrer qu'un certain objet, qu'on pouvait définir, et dont on savait qu'il avait une réalisation, n'en avait qu'une seule. C'était un problème technique, très ardu, difficile à attaquer directement parce qu'on avait vite épuisé tous les moyens dont on disposait. A force de parcourir un champ d'exploration voisin mais disjoint, où les objets d'étude étaient plus nombreux et plus faciles à saisir, je parvenais à une expertise, à une intuition qui pouvait s'appliquer au premier problème. Il s'agissait donc d'un cadre, d'un champ d'exploration indirecte.

JPC : Tu utilisais donc un cadre plutôt que des objets de pensée.

AC : Exactement. Un cadre dans lequel ma pensée pouvait se déplacer, évoluer, alors que dans le contexte trop précis du problème, au pied du mur, la pensée se figeait, bloquée par la difficulté.

JPC : Tu élargissais en quelque sorte le contexte pour laisser apparaître une *variabilité*. Nous voilà en pleine évolution darwinienne. Tu définis une période pendant laquelle s'engendrent des variations conscientes ou inconscientes, pendant laquelle des objets de pensée s'associent, des « cadres » se forment, au sein d'un cadre plus général que celui du problème posé.

AC : Disons pour simplifier que les mathématiciens qui n'arrivent pas à résoudre un problème, ont l'habitude de le généraliser afin de pouvoir en résoudre un cas particulier. Une ouverture permet ensuite d'atteindre un petit morceau du puzzle. On espère, bien entendu, que la résolution de ce cas particulier du problème généralisé, qui a peu à voir avec le problème de départ, permettra de donner une idée qui, elle, s'adaptera. On essaie donc de généraliser pour découvrir plusieurs aspects du problème. On procède ensuite par échelons, relativement petits, pour accéder à l'objet en question.

JPC : On s'efforce donc de combiner ces éléments en les rassemblant dans un cadre élargi, alors qu'auparavant ils étaient disjoints.

AC : Il ne faut pas occulter la distinction entre le processus subconscient de la deuxième étape et le type de programme dont je viens de parler, qu'on explicite et qui appartient au patrimoine culturel. C'est une stratégie qui n'est en rien inconsciente, puisqu'elle est connue de tous les spécialistes. Mais j'ai eu souvent l'impression que le mécanisme cérébral comportait un système qui n'est pas, comment dire, à découvert, qu'on ne perçoit pas directement, mais qui repose sur des mécanismes très analogues.

JPC : Il peut très bien exister un recrutement d'objets de mémoire qui affleure tout juste à ce qu'on peut appeler la *conscience*. Une sorte de travail mental se produit, sans que toutes les opérations qu'il engage soient parfaitement maîtrisées par la volonté. C'est vrai pour les mathématiques comme pour la pensée en général. Une pensée sans langage ¹ est possible. L'expérience de mathématicien que tu rapportes confirme, à mes yeux, l'occurrence d'une phase d'incubation, et suggère que, pendant cette période, des variations darwiniennes qui se recombinent dans le temps sont produites de manière transitoire. A un moment donné, l'une d'elles se trouve *adéquate* au problème posé, et dans un cadre élargi apporte une solution : c'est l'illumination !

Est-ce que tu peux préciser, parce que c'est un point important, quelles sont les conditions de cette adéquation ? Des quantités d'objets mathématiques passagers, conscients ou inconscients, sont évoqués, et puis d'un seul coup, tout s'enclenche, la clé entre dans la serrure et ouvre la porte. Suivant le schéma du darwinisme mental ², après une étape de variations d'un « générateur de diversité », interviennent des « processus de sélection ».

AC : Il est difficile d'assurer qu'au cours de la deuxième phase, un tel rôle est tenu par un générateur de diversité. Ton modèle correspondrait à celui des ordinateurs qui jouent aux échecs. Ceux-ci ont un comportement relativement darwinien : une quantité considérable d'essais sont accomplis, qui n'aboutiraient à rien sans une fonction de sélection qui mesure à la fois le gain obtenu en jouant plusieurs coups et la force de la position qu'ils entraînent. On doit donc introduire une quantité qui exprime ces gains ou la force de la position, et que l'ordinateur, lui, doit pouvoir opti-

1. L. Weiskrantz, *Thought without language*, Fond. Fyssen, Clarendon Oxford, 1988.

2. J.-P. Changeux et S. Dehaene, *Neural models of cognitive functions*, Cognition, 1989 (sous presse).

miser. Pour trouver un mécanisme darwinien dans le fonctionnement du cerveau du mathématicien au travail, il faut d'abord chercher l'analogie de cette fonction de sélection.

Les mathématiciens savent bien que comprendre un théorème ne signifie pas comprendre pas à pas une démonstration dont la lecture peut durer plusieurs heures. C'est au contraire voir la totalité de cette démonstration en un temps extrêmement bref. Le cerveau doit être capable de « vérifier », j'ignore comment, cette démonstration en l'espace d'une ou deux secondes. On est certain d'avoir compris un théorème si l'on a ce sentiment-là. Pas si l'on est capable de parcourir la démonstration sans trouver d'erreur, ce qui ne donne qu'une compréhension locale. Au moment de l'illumination se produit un mécanisme, que je ne saurais définir, qui assure que la clé ouvre bien la serrure. Pour admettre l'existence d'un mécanisme darwinien dans le cerveau, il faudrait comprendre quel type de fonction d'évaluation entre en jeu, au cours de la période d'incubation, pour sélectionner la solution du problème. On pourrait alors très schématiquement dire que la première étape consiste à construire, de manière consciente, une fonction d'évaluation liée à l'affectivité qu'on pourrait exprimer crûment sous la forme : « ce problème-là, je veux le résoudre ». Le mécanisme darwinien correspondrait à l'incubation, l'illumination ne se produisant que lorsque la valeur de la fonction d'évaluation est assez grande pour déclencher la réaction affective.

JPC : Ce n'est pas une sonnette d'alarme, mais une sonnette de plaisir qui signale...

AC : ...que ce qui a été trouvé marche, est cohérent et, pourrait-on dire, esthétique. Ce plaisir, j'en suis certain, est analogue à celui des artistes, lorsqu'ils trouvent une solution, lorsqu'un tableau est parfaitement cohérent et harmonieux. Le fonctionnement du cerveau doit être le même. Mais le mot « darwinien » semble indiquer quelque chose

de caché qui pose le problème de la fonction sélective, de la quantité à optimiser.

JPC : Bien entendu. Mais rien n'est caché. La sélection participe du mécanisme. La réflexion de type darwinien a pour intérêt majeur de définir des étapes qui autrement ne seraient pas distinguées ou resteraient un peu confuses. Un modèle est intéressant s'il conduit à un progrès, sinon dans la compréhension, du moins dans l'analyse.

2. Le cerveau et ses multiples niveaux d'organisation

JPC : Nous voilà prêts à passer à une autre question. Quel peut être le rôle des neurosciences dans la compréhension du mécanisme de production et de traitement des objets mathématiques? Je mentionnerai à nouveau Desanti. Une épistémologie matérialiste forte doit inclure la description de l'appareil de connaissance et de son fonctionnement, c'est-à-dire de notre cerveau et de la manière dont il produit les objets mathématiques. L'effort pour comprendre les bases neurales des mathématiques a donc une importance scientifique fondamentale. Les psychologues « fonctionnalistes », comme Fodor ¹ ou Johnson-Laird, rejettent cette approche qu'ils tiennent pour inutile. Selon eux, il suffit de décrire les processus de la pensée sous forme d'algorithmes. Ils distinguent ce que les auteurs anglo-saxons appellent *mind* (« esprit » en français, mais dans un sens qui exclut toute métaphysique) c'est-à-dire les *fonctions* du cerveau, de son organisation neurale. Structure et fonction sont nettement

1. J. Fodor, *The language of thought*, The Harvester Press, 1976. P.N. Johnson-Laird, *Mental models*, Cambridge University Press, 1983.

distinguées. Mais, décrire les fonctions cérébrales sous forme mathématique a, pour eux, valeur d'explication et suffit pour comprendre l'ensemble du processus. En tant que neurobiologiste, je me suis toujours opposé à cette attitude. Je suis convaincu, au contraire, que tenter de décrire les bases neurales des fonctions cérébrales, et en particulier de celles qu'engagent les mathématiques, permettra d'améliorer la connaissance des mathématiques elles-mêmes.

AC : C'est certain.

JPC : Avant d'en venir aux bases neurales des mathématiques, il me paraît nécessaire de définir la notion de *niveau d'organisation* (voir figure 19). Le travail du biologiste consiste en général à mettre en relation une fonction avec une organisation structurale définie : il s'agit d'établir une relation *causale* entre structure et fonction. Si, avant même d'aborder une recherche, on ne réfléchit pas à cette mise en relation de la fonction avec la structure, on risque fort de faire des erreurs magistrales. Tu en connais déjà certaines. L'une des plus célèbres fut, au XIX^e siècle, la conviction qu'avaient certains biologistes « physicalistes » que la génération spontanée devait exister. Le débat portait initialement sur la nécessité de la présence de levure pour que la fermentation ait lieu. La fermentation semblait une « décomposition chimique » et devait pouvoir se produire intégralement *in vitro*. Ce qui est vrai, Buchner le montrera ! Mais on en déduisait la possibilité de reconstituer une cellule vivante à partir d'une population de molécules en solution, donc de réaliser la génération spontanée. Ce que niait Pasteur à l'époque, et avec raison ¹. D'où venait l'erreur ? Certainement pas de l'existence de « forces vitales », irréductibles aux lois de la physique et de la chimie, qui empêcheraient cette reconstitution ! L'extrême complexité de l'organisation cellulaire, qu'il est encore impossible de

1. C. Debru, *L'Esprit des protéines*, Hermann, Paris, 1983.

reconstituer dans sa totalité, même dans le cas d'organismes aussi simples que les bactéries en levure, n'était pas appréciée à sa juste valeur. Une cellule ne se compose que de molécules. Mais elles forment un ensemble très organisé qui se divise, se multiplie, du fait d'interactions très particulières et très intriquées les unes avec les autres que nous sommes encore très loin de comprendre dans leur intégralité. Les auteurs de l'époque n'établissaient pas une relation correcte entre structure et fonction. Ils suggéraient une mise en correspondance inadéquate entre structure et fonction dépendant de niveaux d'organisation différents.

Lorsqu'on aborde le problème de la relation des mathématiques avec le cerveau, il faut éviter de tomber dans ce travers. Le simple fait de soulever ce problème suscite l'irritation. On vous assène que les mathématiques constituent un monde tellement distinct des neurones, des synapses, etc., qui composent le cerveau, que c'est perdre son temps que de tenter une telle mise en relation. Ce projet rencontre donc une opposition très rigide. Pour que la mise en relation causale d'une structure statique avec une fonction par nature dynamique ait un sens, il faut qu'elle soit effectuée au niveau d'organisation pertinent. Le biologiste doit donc au préalable définir des niveaux hiérarchiques pertinents sur le plan fonctionnel, et ce avant même d'aborder une recherche expérimentale.

Or les philosophes de « l'esprit », en particulier les plus grands, comme Kant, se sont intéressés à cette question. Kant distingue ainsi trois niveaux, qui me paraissent intéressants à rappeler. Celui de la *sensibilité*, définie par la capacité à recevoir des « impressions » par les organes de sens. Celui de *l'entendement*, ou faculté des concepts qui permet la synthèse des éléments sensibles. Celui de la *raison*, qui contient les principes de l'usage des concepts spontanément produits par l'entendement. Ces distinctions kantienues peuvent nous permettre de concevoir trois niveaux

d'abstraction : 1) l'élaboration de représentations à partir des objets du monde extérieur; 2) leur abstraction en concepts; puis 3) l'organisation de ces concepts en abstractions d'ordre plus élevé... tout cela bien entendu *dans* le cerveau. Après les avoir définies, on peut tenter, à ses risques et périls, de mettre en relation ces « facultés » avec des organisations connexionnelles de notre encéphale.

Il est remarquable que, d'une manière indépendante, des informaticiens comme Newell ¹ et Simon ² se soient intéressés aux niveaux hiérarchiques à l'intérieur des ordinateurs.

AC : Cette question est décisive. La comparaison avec les ordinateurs peut permettre, j'en suis persuadé, d'affiner la définition des différents niveaux de l'activité du cerveau en considérant seulement l'activité mathématique.

JPC : Newell et Simon ont défini ce qu'ils appellent un « niveau de la connaissance » (knowledge level), qu'ils situent, pour des ordinateurs théoriques qui n'ont pas encore été réalisés, au-dessus du « niveau symbolique » des ordinateurs habituels. Le niveau de la connaissance s'enrichit en permanence avec l'expérience de nouvelles actions, moyennant un principe, le « principe de rationalité », qui est le suivant : « Si un agent a la connaissance que l'une de ses actions conduira à l'un de ses buts, alors l'agent sélectionne cette action. »

AC : J'aurais à faire une petite critique de cette phrase.

JPC : Tu peux même la redéfinir ! Es-tu d'accord avec ce clivage entre au moins deux niveaux, qui se rapprocheraient l'un de l'entendement (en gros le symbolique), et l'autre de la raison (la connaissance) ?

AC : Je pourrais même définir assez précisément trois niveaux dans l'activité mathématique. Mais je vois plus

1. A. Newell, *The Knowledge level*, Artificial Intelligence, n° 18, 1982, 87-127.
2. H.A. Simon, *The Sciences of the Artificial*, MIT Press, Cambridge, 1984.

difficilement comment les relier à ceux de Kant. J'utiliserai une terminologie différente.

JPC : Bien sûr. J'essaierai ensuite de faire cadrer ces niveaux avec les données des neurosciences.

AC : Je crois que le premier niveau, pour la seule activité mathématique, correspond aux ordinateurs actuels. Des mécanismes préétablis permettent de donner une réponse précise à un problème de type donné, en général de nature calculatoire. Il s'agit par exemple de savoir faire une division, sans comprendre pour autant le mécanisme de ses opérations. Bien sûr, aujourd'hui, les ordinateurs font bien mieux. Mais même pour des opérations très sophistiquées comme calculer des intégrales ou tracer le graphe d'une fonction, le mécanisme est toujours donné à l'avance.

JPC : Nous sommes au niveau des opérations élémentaires.

AC : C'est le niveau du calcul, pas nécessairement des opérations élémentaires. Les opérations peuvent être plus compliquées, mais peu importe. Seul compte le fait qu'accomplir ces opérations n'a pas de conséquences pratiques sur la manière dont elles sont accomplies. Une fois l'addition ou la multiplication apprises, à moins de passer à un autre niveau, on ne change pas de méthode, on l'applique sans comprendre pourquoi. Quantité de gens font des divisions, et parfois pendant longtemps, sans savoir comment. Par automatisme. Les ordinateurs actuels n'ont donc pas dépassé le simple calcul, puisqu'ils n'ont aucune compréhension du mécanisme qu'ils appliquent. Ils appliquent des recettes, qui donnent des résultats et beaucoup plus vite que le cerveau humain, mais qui restent des recettes. Je n'ai jamais été impressionné par un calculateur prodige qui applique des recettes connues, ou par les gens qui ont un raisonnement rationnel parfait, ou qui, pour prendre une image, ne manquent jamais de relever une faute de frappe ou de syntaxe. Pourquoi ? Parce qu'ils en restent au premier niveau.

celui des calculs, qui exclut la compréhension globale du système. Il n'y a pas alors interaction entre le système et les calculs qu'il fait.

Le deuxième niveau est plus difficile à définir...

JPC : Mais est-ce que deux niveaux ne suffisent pas ?

AC : Non. L'activité mathématique comporte vraiment trois niveaux. Mais je ne prétends absolument pas qu'ils correspondent à ceux que Kant décrivait. Je voudrais souligner toutefois la richesse du premier niveau. Il regroupe par exemple l'activité qu'on a en maths sup., qui consiste à tracer des graphes de courbes ou à faire des calculs de cinématique...

JPC : Ce sont les mathématiques bêtes !

AC : J'avais un prof de Taupe, je me souviens, qui me disait souvent : « Je voudrais absolument que les choses faciles, vous appreniez à les faire vite et bien. » Ce ne sont que des recettes qu'on applique. Le deuxième niveau commence lorsqu'il y a interaction entre les calculs effectués et une problématique personnelle. Supposons par exemple qu'on ait deux méthodes pour faire un calcul et qu'on obtienne deux résultats différents. On se trouve alors au deuxième niveau, parce qu'on est obligé de se poser une question, soit sur la validité de la méthode, soit sur des erreurs possibles faites pendant le calcul, soit sur la signification des calculs que l'on effectue. On doit alors tester la méthode, et donc comprendre son but et son mécanisme. Il est bien évident que les ordinateurs, pour le moment, en sont incapables.

JPC : Mais on peut demander à un ordinateur de vérifier sa propre méthode...

AC : Pour le moment, si l'on compare, comme dans la navette spatiale, les résultats obtenus à un même calcul par plusieurs ordinateurs, on élimine les erreurs dues à leur dérèglement. C'est encore très éloigné d'une réflexion de

l'ordinateur sur le but qu'il doit atteindre, ou de la possibilité de changer de stratégie.

JPC : Ce n'est pas tout à fait le niveau de la raison...

AC : Je ne prétends pas que ce soit le niveau de la raison. C'est ce que je qualifierais de deuxième niveau. Lorsque le calcul ne marche pas, ou lorsqu'on trouve deux résultats différents, au lieu simplement d'appliquer une recette et de vérifier si l'on s'est trompé, on change de stratégie pour s'adapter. Imaginons quelqu'un qui, à force de faire des multiplications, trouverait une méthode plus simple pour obtenir le résultat. Ou un ordinateur qui, au jeu d'échecs, parviendrait à comprendre ses erreurs pour ne plus les faire ensuite, ou qui inventerait une stratégie. Au lieu d'avoir en mémoire une liste d'ouvertures, il inventerait une nouvelle ouverture.

JPC : On approche du niveau de la connaissance avec cet ordinateur capable de capturer des erreurs et de proposer une stratégie alternative.

AC : Il est très important qu'il ait une donnée... On retrouve alors la frustration dont je parlais... Ce sentiment, il serait important qu'il soit vécu par l'ordinateur lorsqu'il se trompe, perd aux échecs, ou lorsque sa stratégie n'est pas optimale. Il faudrait qu'il soit stressé, ou inversement qu'il éprouve du plaisir à avoir trouvé une méthode plus efficace, plus rapide. Ce mécanisme subtil ne me paraît pas impossible à réaliser, par exemple pour ce qui est d'une plus grande rapidité dans les calculs. L'ordinateur devrait inventer, améliorer les mécanismes de calcul. C'est peut-être possible dans certains cas, mais il est encore trop tôt.

JPC : Il y aurait en quelque sorte une ré-entrée.

AC : Exactement. Il faudrait que l'ordinateur soit lui-même capable d'améliorer son programme, ce que le cerveau, évidemment, peut faire. Mais les ordinateurs actuels en sont assez loin. Parce qu'on a du mal à définir, de manière autonome, des quantités qui engendreraient frus-

tration ou plaisir chez l'ordinateur, et qui lui permettraient de se débrouiller par lui-même. Cette difficulté m'invite à regarder avec sympathie ce qui, dans le cerveau, lui permet d'avoir des sentiments. Ces derniers jouent un rôle essentiel pour passer au deuxième niveau. C'est analogue à la capacité de construire des hiérarchies de valeurs, de les utiliser et de les modifier. Je connais des mathématiciens qui ont un cerveau purement rationnel, que je qualifierais de premier niveau. Ils me frappent par leur absence de hiérarchie. Ils sont incapables de savoir si un objet d'étude ou un théorème est plus intéressant qu'un autre. Pour autant que leur démonstration est correcte, tous les théorèmes, pour eux, se valent. Le deuxième niveau, au contraire, suppose la possibilité d'apprécier la qualité ou la valeur d'un théorème.

Mais venons-en au troisième niveau, celui de la découverte. A ce niveau, on n'est pas seulement capable de résoudre un problème posé. Mais on peut aussi découvrir – je ne dis pas inventer, parce que ce ne serait pas conforme à la philosophie que j'ai de la préexistence du monde des mathématiques à l'intervention de l'individu – une partie des mathématiques à laquelle les connaissances acquises ne donnent pas un accès direct. On parvient à poser des problèmes nouveaux, à ouvrir des voies inaccessibles auparavant, et à découvrir une partie encore inexplorée de la géographie des mathématiques. Chez le mathématicien, on peut distinguer deux types d'activité. L'une consiste à résoudre des problèmes déjà posés. Et l'autre, à l'occasion d'un problème déjà posé ou d'une réflexion, à créer des outils de pensée, qui n'existaient pas dans le corpus établi et qui permettent de dévoiler une partie encore inexplorée de la réalité mathématique.

JPC : Pour en revenir à Kant...

AC : Je ne dis pas du tout que c'est analogue. Je n'y ai pas réfléchi.

JPC : De toute façon, on n'arrivera jamais à entrer

exactement dans les catégories, telles que Kant les définissait. Peu importe. Il me paraît toutefois essentiel de définir des niveaux de fonction. Je pense que le premier niveau est proche de celui de l'entendement. Quant aux deux autres, je les identifierais à la raison, mais en les hiérarchisant. Notre collègue Gilles-Gaston Granger, qui enseigne la philosophie au Collège de France, a distingué lui aussi deux aspects de la raison ¹. D'une part les aspects *tactiques*, « d'observance d'un lien de principe à conséquence, de calcul propositionnel ou logique ». Et d'autre part les aspects *stratégiques*, « qui conduisent à la détermination d'un champ dans lequel pourra jouer la logique et qui porte sur la plausibilité des buts ou des fins ». La raison tactique ne consiste pas simplement à effectuer des opérations, mais à mettre à l'épreuve le calcul logique, la validité d'une proposition logique.

AC : La vérification de la validité du raisonnement dans un théorème relève, selon moi, du premier niveau. Je ne désespère pas que les ordinateurs en soient bientôt capables.

JPC : Si j'ai bien compris Granger, la raison tactique inclut la possibilité de changer de tactique. Il ne s'agit pas seulement de mettre à l'épreuve une tactique déterminée, mais également d'en élaborer de nouvelles. Autrement le mot « raison » ne se justifierait pas. Cette raison tactique semble proche de ce que tu considères comme le deuxième niveau. Au contraire, élaborer une nouvelle stratégie se rapproche plutôt, me semble-t-il, de la création pure, de l'ouverture d'un nouveau champ de connaissance et d'investigation, de la définition d'une nouvelle catégorie de problèmes.

AC : Oui et non. Il me semble que ce que j'appelais deuxième niveau tient à la fois de la raison tactique et de

1. G.-G. Granger, *Les deux niveaux de la rationalité*, in *Dialectica* n° 39, 1985, p. 355-363.

la raison stratégique de Granger. Ses divisions et les miennes ne se recouvrent pas exactement. Peu importe d'ailleurs. Je parle, moi, non en philosophe, mais en praticien des mathématiques.

3. *Le niveau cellulaire*

JPC : Il ne me paraît pas nécessaire de faire cadrer exactement toutes ces définitions les unes avec les autres. Ce serait une attitude trop réductrice et intransigeante, que je récusé. D'ailleurs, il serait absurde de penser que le cerveau est divisé en compartiments étanches. Néanmoins, la concordance paraît bonne. Mais il est bon de souligner que cette subdivision en niveaux est une opération effective de notre cerveau de scientifiques !

Nous sommes prêts pour aborder les neurosciences. Certains niveaux sont très faciles à distinguer, d'autres plus difficiles. Le plus simple, c'est celui de la cellule nerveuse, du neurone (*voir figure 17*), avec ses prolongements dendritiques qui collectent les signaux vers le corps cellulaire, et l'axone qui propage l'influx nerveux à partir de celui-ci. Comme tu le sais, l'ensemble de notre encéphale se compose d'environ cent milliards de neurones, ce qui est, quand même, un chiffre élevé ! Ces neurones sont reliés entre eux par des zones de contact discontinu ou synapses. Il y en a en moyenne environ dix mille par cellule nerveuse. Soit un nombre total de synapses, dans notre cerveau, de l'ordre de 10^{15} . C'est un chiffre astronomique. Qu'est-ce que cela t'évoque comme niveau de complexité ?

AC : C'est effectivement colossal. On pense au nombre d'Avogadro, assez proche (10^{23}).

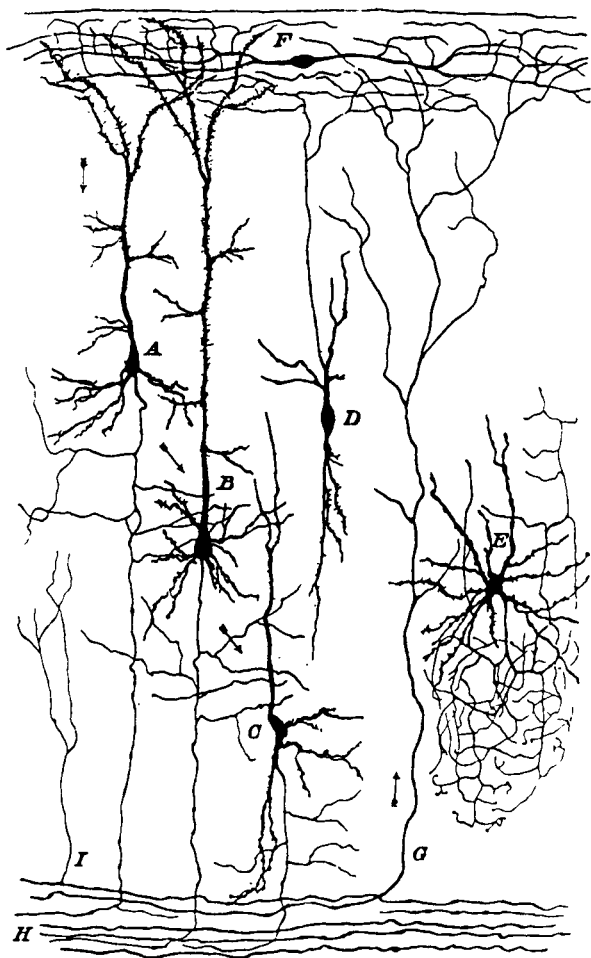


FIGURE 17

Principaux types de neurones du cortex cérébral des mammifères. Les cellules ont été imprégnées à l'argent et paraissent en noir opaque. Les arborisations dendritiques, qui collectent les signaux nerveux, se reconnaissent par de fines épines; l'axone est unique, mais présente des ramifications collatérales souvent perpendiculaires à celui-ci. Les flèches indiquent le sens de propagation des signaux nerveux. A, B, et C: cellules pyramidales; D: cellule à axone ascendant; E: cellule granulaire. (D'après S. Ramon y Cajal, in Cajal on the cerebral cortex J. de Felipe and E.G. Jones, Oxford University Press, 1988.

JPC : Le neurone est la brique élémentaire, ou plutôt la tesselle, pour employer le terme de la mosaïque. Sa fonction est relativement aisée à définir : produire des influx nerveux. Il s'agit d'impulsions électriques dont l'amplitude est d'environ 100 millivolts et la durée de l'ordre de la milliseconde. Ceux-ci se propagent d'un neurone à l'autre, le long de l'axone, à des vitesses inférieures à la vitesse du son, de quelques mètres à quelques dizaines de mètres par seconde, sur des distances qui sont parfois grandes. Dans le cerveau, les axones peuvent atteindre le centimètre, voire la dizaine de centimètres de longueur. Dans l'organisme, l'axone peut atteindre le mètre, puisque les neurones moteurs de la moelle épinière commandent les mouvements des orteils. Ces signaux sont discrets, de l'ordre du tout-ou-rien ; ils se propagent comme des solitons et véhiculent l'essentiel de l'information susceptible d'être transmise par le système nerveux central. Ce sont les « grains d'activité » discrets, universels...

AC : Est-ce que tu ne peux pas garder aussi la composante chimique et hormonale ?

JPC : Bien entendu, puisque c'est sur ce thème que nous travaillons depuis vingt ans. Cette composante chimique est essentielle à la transmission des signaux à travers les connexions et à la régulation d'efficacité des connexions. Une conception du système nerveux où n'interviendraient que des potentiels d'actions électriques dans un réseau de câbles serait trop rigide.

AC : Elle serait un peu trop réductrice.

JPC : Les impulsions électriques qui circulent dans notre système nerveux sont toutes de même nature. Elles sont identiques chez le calmar, la drosophile et l'homme. Elles peuvent être décrites par la même équation de Hodgkin et Huxley. Elles peuvent être engendrées de manière autonome, spontanée, par la cellule nerveuse, en l'absence d'interaction avec le monde extérieur. C'est le cas, en par-

ticulier, lors du rêve. Mais cette genèse d'influx nerveux peut également être provoquée par le contact avec l'environnement. Le système visuel est sujet à ces deux types d'activité. La lumière, après avoir touché les cellules réceptrices de la rétine, entraîne l'activité électrique des cellules ganglionnaires dont les axones forment le nerf optique. Les impulsions électriques cheminent le long du nerf optique, atteignent le corps genouillé latéral, où ils stimulent des neurones relais qui envoient des signaux au cortex cérébral, au niveau de l'aire visuelle. Il y a alors activité évoquée. Mais notre système visuel est également capable d'avoir une activité spontanée, en particulier chez le fœtus, où elle sert vraisemblablement à régler la maturation du système. Il y a donc à la fois genèse spontanée et genèse évoquée d'impulsions nerveuses qui sont, bien entendu, indistinctes une fois produites.

Au niveau du système, le « code » de signalisation associé à ces ondes propagées est extrêmement modeste. En plus des ondes solitaires, on note des trains d'impulsions qui sont soit très réguliers, soit espacés dans le temps de manière définie, par exemple exponentielle. Il peut exister des rafales périodiques d'impulsion... avec une régularité d'horloge. Mais ce n'est pas un code « morse » qui détiendrait un langage. En fait, la « sémantique » se trouve au niveau de l'anatomie des connexions. L'état d'activité définit le recrutement d'un ensemble particulier de neurones au sein d'un réseau beaucoup plus complexe. Une sorte de « contraste » va apparaître entre neurones actifs et neurones inactifs, ou neurones plus actifs que d'autres, ou entre neurones aux activités corrélées entre elles ou non. Tout ce que je viens de décrire se situe au niveau cellulaire.

AC : Avant qu'on aille plus loin, une question d'ordre général. Ce qui me frappe, si on ne regarde que les connexions électriques, c'est que le courant se propage à une vitesse beaucoup plus faible que la vitesse de la lumière, un petit

peu comme un soliton. Sa vitesse de propagation est de l'ordre de celle du son.

JPC : Et même inférieure.

AC : Ce phénomène est pour moi un mystère. J'aimerais bien un jour comprendre pourquoi ce mécanisme peut avoir un effet positif. En effet, la découverte de la supra-conductivité à température relativement haute crée l'espoir de fabriquer des ordinateurs qui fonctionnent mille fois plus vite que ceux d'aujourd'hui, grâce à l'amélioration du système de propagation. J'espère qu'on expliquera un jour pourquoi la petite vitesse de propagation de l'information dans le cerveau et son analogie avec le soliton ont un rôle positif plutôt que négatif.

JPC : On ne peut pas dire que cette vitesse a un rôle positif ou négatif. C'est un fait. Il faut raisonner autrement, et adopter, à nouveau, un point de vue évolutionniste. L'organisation cellulaire, au cours de l'évolution, s'est développée chez les bactéries, puis chez les cellules dites supérieures, à partir des éléments qui étaient alors disponibles. C'est pourquoi François Jacob parle de bricolage¹. A partir de ces éléments, s'est constituée une membrane lipidique imperméable, avec des systèmes de transport sélectifs d'ions Na^+ , K^+ , Ca^{++} ..., puis création d'un gradient électrochimique et enfin genèse d'un potentiel de membrane. Ce potentiel électrique a, ensuite, été « exploité » pour produire un signal qui se propage. Cette règle de propagation a été conservée et utilisée dans des systèmes beaucoup plus complexes.

Il est vraisemblable que l'influx nerveux ou potentiel d'action est apparu chez les êtres unicellulaires très primitifs. Des signaux électriques de type propagé sont enregistrés en effet chez des paramécies ou des algues unicellulaires. Dans les organismes multicellulaires plus complexes

1. F. Jacob, loc. cit., p. 85.

de type « animaux », certaines cellules se différencient en produisant des câbles qui servent à communiquer, à transmettre des ordres à d'autres cellules. Le système nerveux se développe comme centre de contrôle de l'organisme. Les cellules spécialisées qui le composent utilisent les propriétés électriques existantes pour propager des signaux, donner et recevoir des ordres des autres cellules de l'organisme. La « petite vitesse » du transfert des signaux dans le système nerveux est issue de son histoire évolutive. Les organismes vivants primitifs ne devaient pas disposer, pour produire des systèmes de signalisation plus rapide, d'éléments cellulaires permettant d'exploiter les propriétés de supra-conductivité de la matière.

Je voudrais maintenant revenir sur ce que tu as dit de l'importance de la chimie dans la signalisation neuronale. Effectivement, il ne faut pas concevoir le système nerveux comme une machine « rigide », purement électrique. Le transfert d'informations au sein du système nerveux comporte d'importantes possibilités de régulation, qui ont permis l'apprentissage et qui se situent au niveau du neurone lui-même et de la genèse des impulsions, mais aussi, au niveau des synapses, au point d'articulation entre cellules nerveuses. A ce niveau, si les membranes des cellules en présence sont suffisamment rapprochées, l'impulsion électrique peut passer directement d'une cellule à l'autre. Le plus souvent une substance chimique, un neurotransmetteur, prend le relais. Celui-ci s'accumule dans la terminaison nerveuse et l'arrivée de l'influx nerveux déclenche sa libération dans l'espace synaptique. Là, il diffuse rapidement jusqu'à la cellule suivante où il se lie sur des *récepteurs* spécifiques et, à ce niveau, provoque une réponse électrique due à l'ouverture de canaux ioniques. Ces récepteurs, que nous avons beaucoup étudiés dans mon laboratoire, sont la cible d'agents pharmacologiques puissants comme le curare, le LSD, la morphine, le valium et... la nicotine. Ils se trouvent

à un point critique dans le transfert d'information entre cellules nerveuses. C'est pourquoi Thierry Heidmann et moi-même¹ avons proposé un modèle de régulation d'efficacité de la synapse au niveau de ces récepteurs postsynaptiques de neurotransmetteurs (*voir figure 18*). Ces protéines, qui traversent la membrane, sont, en effet, susceptibles d'exister sous plusieurs conformations réversibles possédant des efficacités de réponse différentes. Elles peuvent passer d'un état à l'autre par des transitions moléculaires relativement lentes. Et celles-ci sont susceptibles d'être réglées par des signaux électriques ou chimiques, et même par plusieurs signaux à la fois. Ces récepteurs sont donc capables d'intégrer plusieurs signaux élémentaires, simultanément dans l'espace et le temps. A un point tel, que cette propriété a intéressé des chimistes qui pensent pouvoir introduire des molécules de ce type dans les ordinateurs. Une discipline nouvelle est d'ailleurs en train de naître : la « bionique ». Elle a obtenu peu de résultats pour l'instant. Mais on peut rêver à des transistors, des microcircuits intégrés... composés de molécules de récepteurs de quelques milliardièmes de mètre!

4. Des circuits élémentaires aux objets mentaux

Considérons maintenant le niveau d'organisation immédiatement supérieur, celui des circuits de neurones (*voir figure 19*). Il est évident que les neurones peuvent s'assembler et se spécialiser dans une fonction particulière,

1. T. Heidmann et J.-P. Changeux, *Un modèle moléculaire de régulation d'efficacité d'une synapse chimique au niveau postsynaptique*, C.R. Acad. Sc., 295, Paris, 1982, p. 605-670, et T. Heidmann et J.-P. Changeux, *Allosteric receptors and molecular models of learning*, in « Synaptic functions », G.M. Edelman, W.E. Gall, W.M. Cowan, p. 549-601.

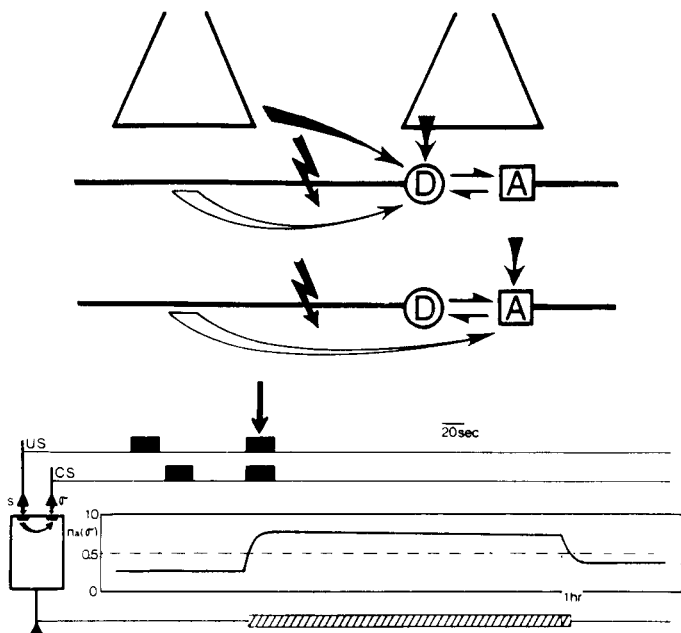


FIGURE 18

Modèle moléculaire de régulation d'efficacité d'une synapse chimique par les transitions de conformation d'un récepteur post-synaptique.

Le schéma du haut représente deux terminaisons nerveuses (triangles) en contact avec la surface d'un même neurone (ligne horizontale) contenant des molécules du récepteur qui reconnaît le neurotransmetteur. Celui-ci est susceptible d'exister sous deux formes interconvertibles, l'une plus efficace dans le transfert d'information (A) que l'autre (D). Des signaux chimiques (neurotransmetteurs, neuropeptides...) et/ou électriques (flèches) produits par la synapse de gauche règlent l'efficacité de la synapse de droite, lorsqu'ils affectent conjointement l'un des deux états du récepteur post-synaptique. Leurs effets relatifs sur A ou D augmentent (ligne inférieure) ou diminuent (ligne supérieure) l'efficacité de la synapse.

Le schéma du bas montre l'évolution dans le temps de l'efficacité d'une synapse conditionnelle (CS) désignée par σ lorsque sa stimulation est effectuée conjointement à celle de la synapse régulatrice non conditionnelle (US) désignée s. La conjonction dans le temps des deux stimulations (flèches) crée un accroissement d'efficacité ($\eta_a(\sigma)$) qui persiste plusieurs minutes (trait hachuré). D'après Heidmann & Changeux, 1982 et Changeux & Heidmann, 1987.

comme les actes réflexes : ceux de la marche, les premières étapes de la vision... La rétine, par exemple, est un réseau fort compliqué qui, à partir des photons capturés par les photorécepteurs, élabore une première représentation du monde extérieur. Un deuxième niveau, pleinement réalisé chez les invertébrés comme le ver de terre ou la limace, donne ce que les éthologues appellent « actes fixes » : picorer, voler, s'accoupler, capturer une proie.

AC : Sans savoir s'il est inné ou acquis, on dit que c'est un deuxième niveau d'organisation.

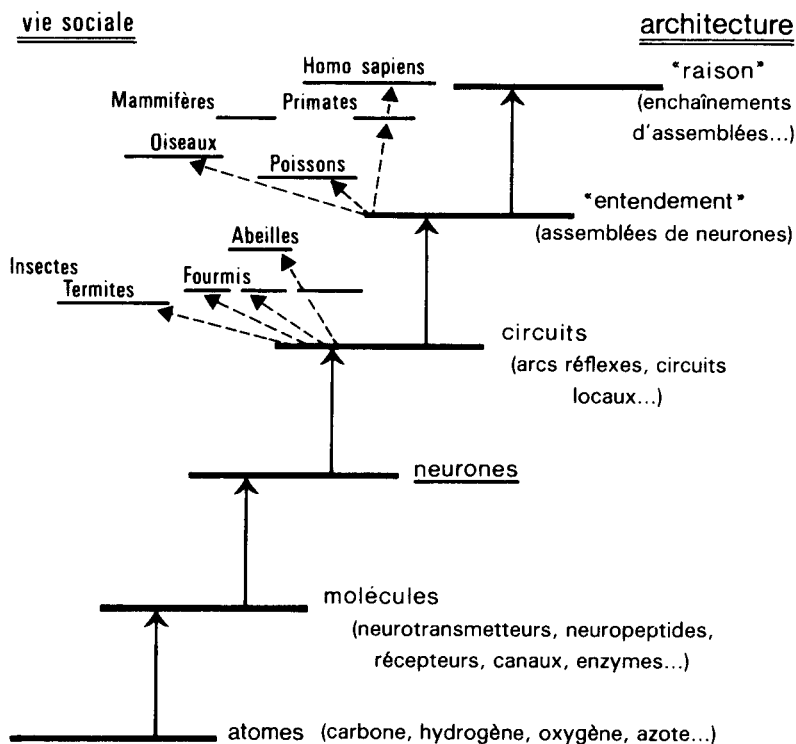


FIGURE 19
Niveaux d'organisation dans le système nerveux.

JPC : C'est celui des *circuits élémentaires* de la moelle épinière, du tronc cérébral, des « mini-cerveaux ». De petits ensembles de neurones, chez les invertébrés, se rassemblent en ganglions, qui s'enchaînent les uns avec les autres, de manière répétitive.

Vient ensuite un autre niveau. Certains invertébrés particulièrement évolués, comme les pieuvres qui ont des comportements assez proches de ceux des vertébrés, mais surtout les vertébrés supérieurs et l'homme construisent des « représentations ». Leur système nerveux possède la faculté d'assembler des cellules nerveuses entre elles pour coder, par exemple, la conduite d'un véhicule dans le cas d'une représentation « motrice ». Mais il existe aussi des représentations de type sensoriel, des représentations plus « abstraites ». Une organisation *hiérarchique* se greffe sur une organisation *parallèle* en multiples cartes neurales dont nous avons déjà parlé.

Les bases neurales du « code » engagé par ces représentations ont été étudiées, très en détail, par Georgopoulos¹ dans le cas du pointage de la main chez le singe éveillé. Celui-ci a enregistré l'activité individuelle de plusieurs centaines de neurones du cortex moteur pendant que le singe pointe sa main dans une direction définie. Il a essayé de définir comment ce programme moteur était codé (ou « représenté ») au niveau de la population de neurones enregistrés. Il a pu montrer que chaque cellule de cette population présente une activité maximale lorsque le singe pointe la main dans une direction particulière ou direction préférée qui signe sa spécificité (*voir figure 20*). Pour chaque neurone, on définit un *vecteur*, dont l'orientation correspond à la direction optimale, et la longueur à l'activité de ce neurone lorsque le singe pointe la main dans une direction définie.

1. A.P. Georgopoulos, A.B. Schwartz et R.E. Kettner, *Neuronal population coding of movement direction*, Science n° 233, 1986, p. 1357-1460.

Cette longueur varie lorsque la direction de la main change. Elle correspond, en quelque sorte, au « vote » de ce neurone particulier, ou à l'impôt payé par ce neurone, au cours du codage de la direction de la main par l'ensemble de la population. La direction vers laquelle le singe pointe sa main est en effet représentée, avec moins de 10 % d'erreur, par la somme vectorielle de ces vecteurs « neuronaux » élémentaires. Le vecteur de direction de pointage de la main coïncide avec la somme vectorielle des « votes » de la population de neurones.

AC : A l'addition ? C'est extraordinaire. Il y a une représentation cartésienne, par coordonnées...!

JPC : Oui, par des activités individuelles. La somme vectorielle de ces activités « microscopiques » correspond de très près à la direction « macroscopique » de pointage de la main du singe. Il y a codage par une *assemblée de neurones*, et je pense que ce modèle est général (*voir figure 21*). A un certain niveau de complexité du système nerveux central, s'organisent des « représentations » ou « objets mentaux »¹ qu'on peut définir à la fois par l'état d'activité des neurones de la population et par le graphe de ces neurones. Chaque neurone de la population diffère de son voisin. Chacun possède une spécificité fonctionnelle au sein de cet ensemble, une individualité, une « singularité ».

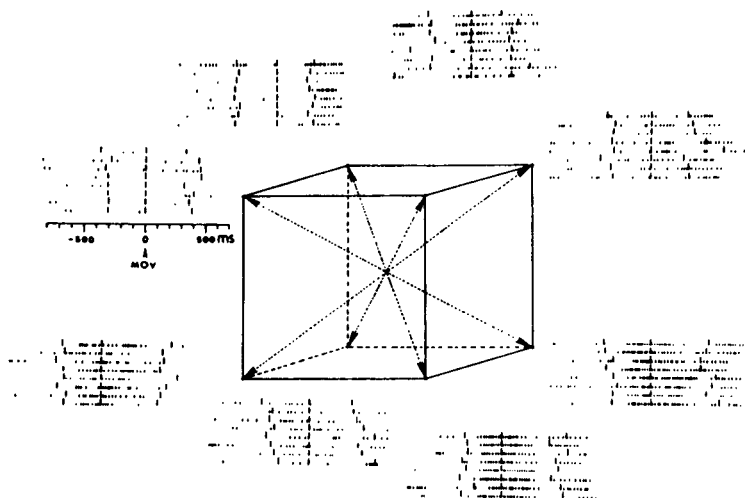
AC : Tu disais qu'il était important que la pieuvre puisse acquérir cela. Elle ne l'avait pas de manière innée.

JPC : Il faut séparer les problèmes. J'essaie pour le moment de définir des niveaux d'organisation et je n'ai pas fini.

Nous arrivons donc au niveau qui pourrait être dit *symbolique*, ou appelé *entendement*. A ce niveau d'organisation, on peut définir en termes physiques des représentations mentales. A un niveau supérieur, que j'appelle le

1. J.-P. Changeux, *L'Homme neuronal*, Fayard, Paris, 1983.

A



B

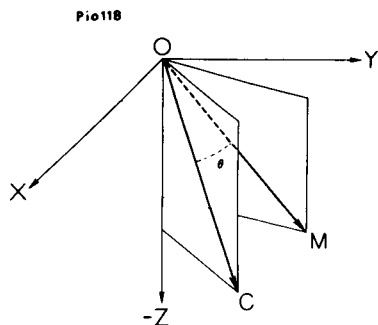


FIGURE 20

États d'activité d'un même neurone du cortex moteur du singe éveillé lorsque celui-ci tente d'atteindre avec sa main une cible située successivement dans 8 directions de l'espace à trois dimensions représentées ici par des flèches. A) L'activité électrique enregistrée est figurée par de petites barres verticales qui correspondent chacune à un influx nerveux. Chaque ligne superposée correspond à un essai différent du singe. La ligne verticale commune à tous les enregistrements (Mov) indique le début du mouvement. L'examen de l'ensemble des enregistrements montre que ce neurone répond de manière optimale (petites barres très rapprochées) lorsque le singe pointe sa main dans une direction privilégiée : ici vers « 4 heures 30 » dans le plan de la figure.

B) L'amplitude de la réponse (la fréquence de décharge) lorsque le singe pointe la main dans une direction donnée (M) est une fonction linéaire du cosinus de l'angle formée entre la direction du mouvement et la direction préférée (C) du neurone. Cette amplitude constitue l'« impôt » payé par le neurone au codage du mouvement par l'ensemble de la population. (D'après Georgopoulos, A., Neural Interpretation of Movement : Role of Motor Cortex in Reaching, 1988, FASEB J., 13, 2849-2857.)

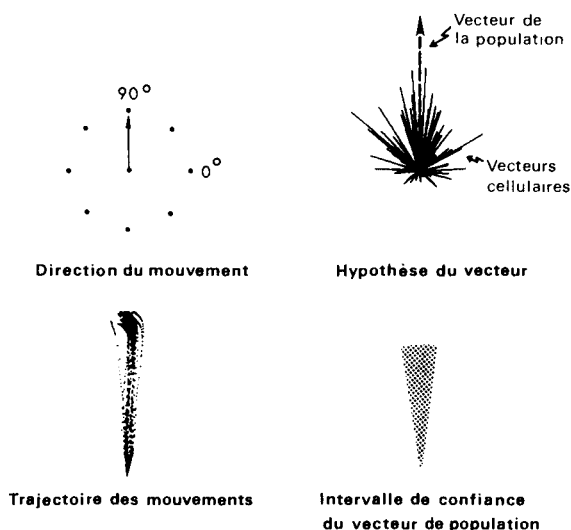


FIGURE 21

Codage par une population de neurones du cortex moteur de la direction du mouvement, ici, vers midi (90°), dans une surface de travail à deux dimensions. Un singe bien entraîné réalise une famille de trajectoires de mouvement assez resserrée. Suivant l'hypothèse du vecteur, la somme des « votes » de chaque neurone enregistré (241 cellules au total) ou vecteurs cellulaires représentés ici par un trait continu, donne un vecteur de population (trait discontinu) qui se trouve orienté dans la direction du mouvement, avec un intervalle de confiance de 99 %. (D'après Georgopoulos, 1988.)

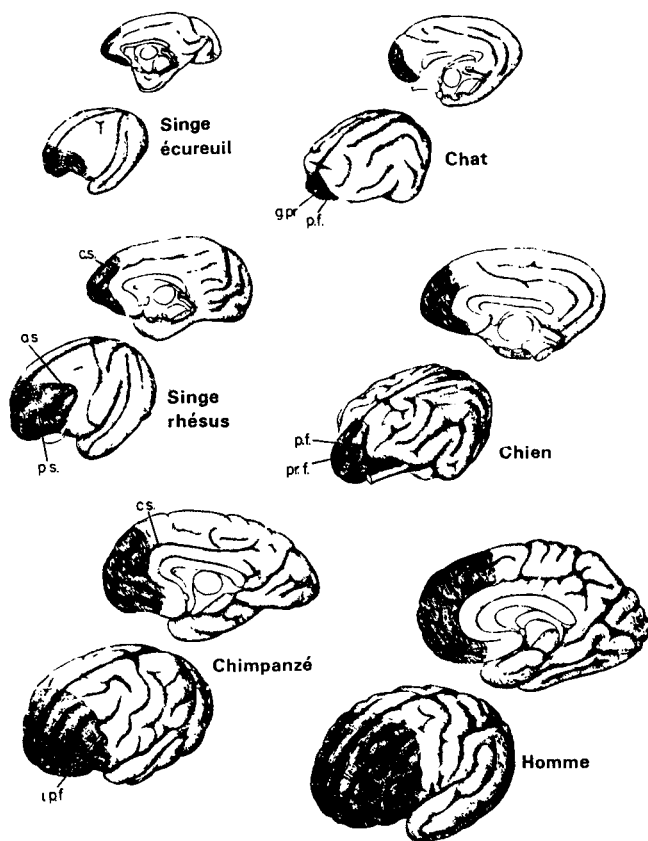


FIGURE 22

Évolution de la surface relative du cortex préfrontal (hachuré) chez les mammifères. Des mammifères primitifs à l'*Homo sapiens*, la surface relative du cortex préfrontal s'accroît. Elle est de 3,5 % chez le chat, 7 % chez le chien, 8,5 % chez le lémur, 11 % chez le macaque, 17 % chez le chimpanzé et 29 % chez l'*Homo sapiens*. Sillons et fissures servent de repères pour délimiter le cortex préfrontal d'une espèce à l'autre : pf = fissure présylvienne, ipf = fissure précentrale inférieure. (D'après Fuster, J., 1980, *The Prefrontal Cortex*, Raven Press, New York.)

niveau de la *raison*, des enchaînements de représentations se forment. Des « assemblées d'assemblées » se développent dans le temps. La temporalité est extrêmement importante et nous n'en avons pas suffisamment parlé.

Dans notre encéphale, le domaine le plus antérieur du cerveau, le cortex frontal (*voir figure 22*), semble être engagé d'une manière définie dans cette fonction. Pour essayer de l'illustrer, je citerai un exemple tiré de travaux de recherche clinique effectués sur des patients souffrant de lésions du lobe frontal. Le plus classique est celui de la réponse au test de Milner et Pétridès ¹. L'examineur demande au sujet de classer un ensemble de cartes suivant une règle définie. Par exemple, l'identité de couleur. Trois cartes sont rouges, la quatrième doit l'être également sinon l'examineur signale l'erreur. On continue avec trois autres cartes, rouges également, le sujet choisit une carte rouge. Il suit toujours la même règle. Soudain, sans prévenir le sujet, l'examineur change de stratégie. La règle portera, par exemple, sur le nombre de figures : seulement des as... Le sujet va d'abord faire des erreurs et continuer à sélectionner des cartes rouges, ce que l'examineur lui signale. Après un certain nombre d'erreurs, le sujet normal se rend compte du changement de stratégie. Le malade qui souffre d'une lésion du cortex frontal, lui, ne s'en rend pas compte. Il persévère dans l'erreur. Suivant Milner et Pétridès, le sujet malade ne formule pas les hypothèses qui lui permettent de progresser dans la réponse au test. Il a perdu une fonction très élémentaire, mais caractéristique, du niveau de la raison.

AC : Cela provient donc d'un endroit précis.

JPC : Oui. Au départ, le sujet présente des troubles neurologiques, par exemple à la suite d'un accident vasculaire. Il consulte un neurologue qui examine son cerveau

1. B. Milner et M. Pétridès, *Behavioural effects of frontal lobe lesions in man* Trends in Neuroscience, 1984, p. 408-414.

par scanographie. Des domaines précis du lobe frontal présentent des lésions. Le malade peut alors être soumis à des tests qui permettent de définir les déficits fonctionnels associés à ces lésions. Inversement, le neurologue peut, en examinant un patient, déceler des réponses anormales à un test particulier et diagnostiquer une lésion frontale, que l'on confirmera par la suite par scanographie. Donc le lobe frontal contribue, en quelque sorte, à ce que j'appelle les *architectures neurales de la raison*¹. Le neuropsychologue anglais Shallice² a bien fait cette distinction. Il sépare ce qu'il appelle les fonctions « de routine » des fonctions de « supervision attentive », qui, selon lui, participent à la capture d'erreurs, à la formulation de nouvelles hypothèses, à l'invention de nouvelles stratégies. On peut donc considérer raisonnablement qu'il existe des domaines de notre encéphale qui participent à l'élaboration de la pensée rationnelle. Fait remarquable, qui appuie cette conclusion, le lobe frontal augmente de surface relative par rapport au reste du néocortex au cours de l'évolution, et s'accroît de manière considérable du rat au singe et du singe à l'homme.

AC : Cela correspond exactement à ma description du deuxième niveau.

JPC : Dans ce cas précis, oui. Mais c'est peut-être aussi le troisième, bien que les tests en question ne le mettent pas suffisamment en évidence.

AC : Il faut conserver une séparation entre le deuxième et le troisième niveau. Tes remarques sur le lobe frontal qui correspond aux fonctions de deuxième niveau m'ont convaincu. Pour le troisième, je n'en sais rien.

JPC : Le lobe frontal possède un rôle important dans la genèse d'hypothèses. Les hypothèses mises en œuvre par

1. J.-P. Changeux, *Molécule et mémoire*, Bedou, 1988.

2. T. Shallice, *Specific impairments in planning*, Phil. Trans. R. Soc. Lond. B. n° 298, 1982, p. 199-209. T. Shallice, *From neuropsychology to mental structures*, Cambridge University Press, 1988.

le jeu de cartes de Milner et Pétridès sont très rudimentaires. Il y a vraisemblablement, à son niveau, genèse d'hypothèses beaucoup plus complexes, mais elles vont être très difficiles à mettre en évidence. Pour les découvrir, il aurait fallu mettre le cerveau d'Archimède dans une caméra à positon quelques fractions de seconde avant qu'il ne crie « Eurêka » !

5. *Neuropsychologie des mathématiques*

JPC : Conformément aux distinctions de niveaux d'organisation dans les fonctions du cerveau, des lésions localisées de l'encéphale permettent de « disséquer » les facultés mathématiques. Hécaen ¹, que tu as peut-être connu, et qui fut un très grand neuropsychologue français, distinguait plusieurs catégories de déficits.

Dans le cas de l'« alexie et/ou agraphie des nombres », le sujet ne lit plus et n'écrit plus les nombres, mais peut conserver l'usage des lettres. Hécaen a pu montrer que l'hémisphère gauche, et plus particulièrement le lobe pariétal de l'hémisphère gauche, intervient dans la lecture et l'écriture des nombres.

Les patients sujets à une « acalculie de type spatial » alignent mal les nombres. Ce déficit porte, semble-t-il, sur un système de pointage visuo-moteur qui permet à la fois de lire et d'ordonner les nombres. Dans ce cas, c'est plutôt l'hémisphère droit qui intervient dans la commande des mouvements de l'œil.

Un troisième déficit s'appelle « anarithmétie ». C'est un déficit du calcul lui-même. Le patient n'arrive plus à faire des calculs, indépendamment du fait qu'il peut les lire, les

1. H. Hécaen et M. Albert, *Human neuropsychology*, Miley, New York, 1978.

écrire et les disposer correctement. Il s'agit dans tous les cas de déficits liés au premier niveau dont nous parlions.

Luria a d'ailleurs fait une subdivision intéressante. Il considère que l'ensemble de ces déficits relève du domaine pariéto-occipital du cortex. Il les distingue de ceux qui affectent le lobe temporal, et entraînent des troubles de la mémoire. Des lésions de ce type font que le sujet ne retient plus ce qu'il est train de faire. Il n'arrive plus à suivre le fil de ses calculs.

Les malades frontaux présentent des troubles d'une autre nature. Ils ne perçoivent plus le problème qu'ils ont à résoudre. Ils perdent le fil, ils n'arrivent pas à avoir un raisonnement suivi, donnent des réponses impulsives, un peu au hasard, et persévèrent dans leur erreur. L'enchaînement de soustractions successives est d'ailleurs utilisé comme test, parmi beaucoup d'autres, pour détecter des lésions frontales. On peut donc considérer que le lobe frontal intervient dans l'enchaînement d'opérations mathématiques, dans la résolution de problèmes et même dans la proposition de problèmes. Il semble donc bien correspondre aux deuxième et troisième niveaux.

AC : Pas vraiment au troisième.

JPC : Les tests qui sont utilisés sont, par nécessité, des tests élémentaires.

AC : Mais on ne peut pas faire de tests pour le troisième niveau.

JPC : Pourquoi ? Il faudrait que tu en imagines un qui soit utilisable par des non-mathématiciens. Un autre test utilisé avec les malades frontaux consiste à leur faire lire puis raconter une histoire. On leur lit par exemple « Le Petit Chaperon Rouge » (Lhermitte) ou « Le Coq d'Or » (Luria), et on leur demande ensuite de reconstituer l'histoire.

AC : On en reste au deuxième niveau dans ce cas.

JPC : Non, des éléments de l'histoire sont reconstitués,

mais l'ensemble est incohérent. La fin arrive avant le début, les épisodes sont mélangés...

AC : C'est de l'ordre de l'organisation, pas encore de l'imagination.

JPC : C'est vrai. Mais trouve-moi donc un test objectif d'imagination. Les neuropsychologues seront ravis.

AC : Je ne sais pas. Mais je voulais te poser une question : on dit très souvent que les mathématiciens perdent leur créativité en vieillissant. C'est un phénomène assez connu. Qu'en penses-tu ?

JPC : Le cortex frontal est sujet à vieillissement relativement rapide, en particulier dans la maladie d'Alzheimer. Les sujets qui en sont atteints peuvent en effet perdre très rapidement la faculté de calcul et la mémoire. Qu'ils perdent leur créativité scientifique, c'est vraisemblable aussi...

AC : Il doit bien être possible de distinguer plus précisément le troisième niveau du deuxième...

JPC : C'est particulièrement difficile dans le cadre des opérations habituelles. Le malade frontal, tel que le décrit Lhermitte ¹, est un malade qui « adhère à l'environnement ». Chaque fois qu'on lui donne un objet, il l'emploie. On le met en présence d'un stylo, il l'utilise. De lunettes, il les met sur son nez. Un marteau, il plante un clou. Il est en prise directe avec le monde extérieur, tout en conservant l'usage des mots. Il effectue normalement les activités de routine, mais il n'arrive pas à résoudre les problèmes qui se présentent dans des situations nouvelles. L'imprévu est un obstacle considérable pour un sujet atteint de certaines lésions frontales.

1. F. Lhermitte, J. Derouesné et J.L. Signoret, *Analyse neuro-psychologique du syndrome frontal*, Revue Neurologique n° 127, 1972, p. 415-440.

6. La transition d'un niveau à l'autre par variation-sélection

JPC : Une nouvelle question nous permettra, je pense, de préciser la distinction entre deuxième et troisième niveaux. Comment passe-t-on de l'un à l'autre? La thèse que j'ai développée depuis plusieurs années ¹, et qui a éga-

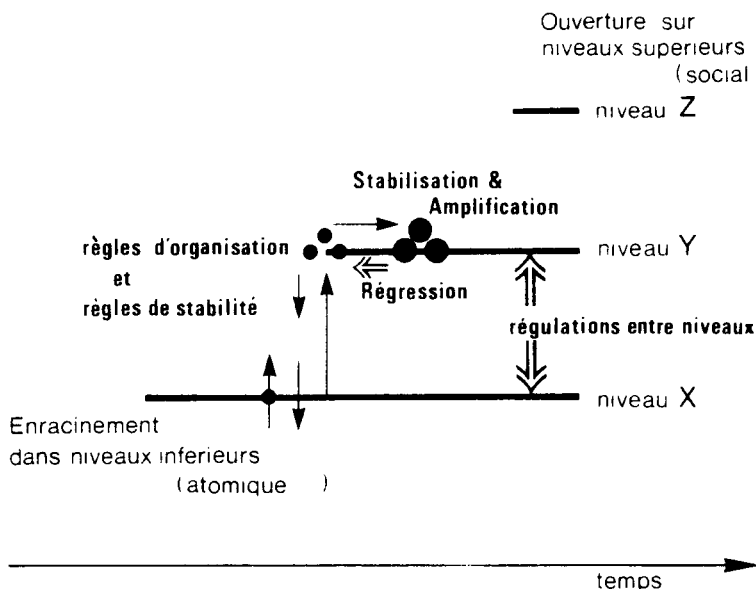


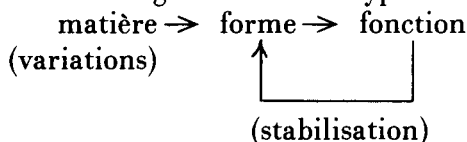
FIGURE 23
Darwinisme généralisé.

1. J.-P. Changeux, *Le cerveau et l'événement*, Communications, 18, 37-47, 1972; J.-P. Changeux, P. Courrèges et A. Danchin. *A theory of the epigenesis of neuronal networks by selective stabilisation of synapses*. Proc. Nat. Acad. Sc., USA n° 70, 1973, p. 2974-2978; J.-P. Changeux et A. Danchin, *Selective Stabilisation of developing Synapses as a Mechanism for the Specification of Neuronal Networks*, Nature, 264, 705-712, 1976; J.-P. Changeux, T. Heidmann & P. Patte, *Learning by Selection*, 115-133, in *The Biology of Learning*, P. Marler & H.S. Terrace ed., Springer Verlag, Berlin, 1984.

lement été évoquée par d'autres auteurs¹, constitue une sorte de « darwinisme généralisé » (voir figure 23) appliqué à la transition d'un niveau à l'autre, quel que soit ce niveau. L'idée est que la transition d'un niveau donné au niveau suivant requiert deux composants fondamentaux : un générateur de diversité et un système de sélection. A un niveau donné, des éléments se recombinent entre eux, varient de manière aléatoire, et élaborent des « formes » transitoires qui relèvent du niveau d'organisation immédiatement supérieur. Ces formes s'engendrent à partir d'éléments déjà structurés, donc pas nécessairement à partir d'atomes. Il y a donc production de variations « darwiniennes » qui peuvent accéder transitoirement à un niveau supérieur d'organisation. Un mécanisme de sélection stabilise ensuite certains de ces états transitoires et ainsi engendre un niveau d'organisation plus élevé.

AC : Quel est ce mécanisme de sélection ?

JPC : Le modèle général est du type :



La fonction agit de manière rétroactive sur la transition « matière-forme ». Le critère de sélection est donc lié à la fonction « nouvelle » déterminée par une forme transitoire produite par le générateur de diversité. Si cette nouvelle fonction correspond à une action sur le monde extérieur qui favorise la survie de l'organisme, elle est sélectionnée.

AC : A l'intérieur du cerveau ou à l'extérieur ?

JPC : J'ai d'abord essayé de te présenter un modèle

1. G. Edelman, *Neural Darwinism*, Basic Books, New York, 1987. G. Edelman, *Group selection and phasic reentrant signalling: a theory of higher brain fonction*, MIT Press, Cambridge Mass., 1978, p. 51-100. N. Jerne, *Antibody and learning: selection versus instruction*, in *The Neuroscience: a study program*, Éd. G.C. Quarton, T. Melnechuck, and F.O. Schmidt, New York Rockefeller Univ., 1967, p. 200-205.

formel très général, que j'espère valable, quel que soit le niveau d'organisation considéré dans l'état initial. Essayons maintenant de l'appliquer. Le cas le plus simple et le plus connu est celui de l'évolution des espèces. Le générateur de diversité se situe au niveau du génome, de l'ADN chromosomique. Les variations « darwiniennes » – mutations, recombinaisons, duplications géniques, transfert de matériel chromosomique – sont des événements aléatoires mais rares, qui entraînent, *secondairement*, des modifications du « phénotype » de l'organisme, qui peuvent s'accompagner d'une « adaptation » à des conditions d'environnement particulières. La ségrégation de combinaisons génétiques particulières peut également se produire, sans qu'une sélection intervienne, à la suite d'un isolement géographique : c'est ce qu'on appelle l'évolution non darwinienne, mais je n'aime pas trop ce terme. Certains éléments « neutres » se maintiennent, tandis que d'autres disparaissent.

Le système nerveux est un organe parmi d'autres. Mais il jouit d'un statut particulier. Les connexions entre cellules nerveuses, les synapses, ne se forment pas en une seule fois, mais à la suite d'un long et complexe processus de développement qui se poursuit, chez l'homme, jusqu'à la puberté. Celui-ci est donc sujet à une évolution *interne* à l'organisme. On distingue en fait deux types, au moins, d'évolution interne : une évolution du *nombre* de connexions, qui a lieu au cours du développement, et une évolution de l'*efficacité* des connexions entre neurones, et donc de leur état d'activité, ce qui est beaucoup moins lourd que le changement de la connectivité.

Considérons d'abord le premier type d'évolution, l'évolution par « épigénèse », au cours du développement embryonnaire et post-natal de l'organisation cérébrale. Au départ, le déterminisme génétique très profond de l'organisation cérébrale implique que le cerveau de l'homme reste celui de l'homme et se distingue de celui du singe. Les gènes

qui interviennent dans ce développement sont en cours d'étude chez les vertébrés. Ils ont fait l'objet d'importantes recherches chez la drosophile, qui, bien que mouche, possède, comme nous, une tête, un thorax, un abdomen et des pattes. Quelques-uns des déterminants génétiques qui fixent les coordonnées cartésiennes de l'embryon (tête-queue, dos-ventre) ont été récemment identifiés¹ : ceux qui règlent la segmentation du corps (le fait que le corps est formé de segments successifs, un peu comme un petit ver), et enfin ceux qui déterminent l'identité d'un segment (céphalique avec antennes et mandibules; thoracique avec ailes, pattes; abdominal avec organes génitaux, etc). Ces trois ensembles de gènes s'expriment de manière différentielle et séquentielle au cours du développement embryonnaire et post-natal. De cette imbrication d'expressions génétiques résulte un organisme qui possède une architecture d'ensemble, un plan d'organisation qui est le même, ou qui est très voisin, d'un individu à l'autre au sein d'une même espèce.

Il est légitime de suggérer que l'expansion du cortex frontal, que l'on suit chez les mammifères, de la souris à l'homme, est soumise au contrôle de quelques-uns de ces gènes. Leur nombre est vraisemblablement peu élevé. L'ADN du chimpanzé, en effet, présente 99 % d'identité avec l'ADN de l'homme. On peut concevoir que si quelques-uns de ces gènes restent actifs un peu plus longtemps dans la partie antérieure de l'ébauche embryonnaire de l'encéphale, il s'ensuivra un accroissement différentiel de la surface du cortex frontal. L'organisation d'ensemble de notre encéphale, l'essentiel de notre architecture cérébrale, est soumis au pouvoir des gènes.

Néanmoins, le pouvoir des gènes a des limites. Comment

1. C. Nüsslein-Volhard, H.G. Frohnhöffer, R. Lehman, *Determination of anteroposterior Polarity in Drosophila*, Science, 238, 1675-1681, 1987; W. Gehring, *Homeotic Genes, the Homeobox of the genetic control of development*, Cold Spring Harbor Symp. Quant. Biol, 50, 243-251, 1985.

les mettre en évidence? On peut d'abord comparer la connectivité d'un même neurone identifiable par sa forme et sa position dans deux individus génétiquement identiques, chez deux *vrais* jumeaux. L'expérience a été réalisée par les Levinthal¹ chez un crustacé parthénogénétique, la daphnie, une puce d'eau qui possède un système nerveux simplifié composé d'un nombre fixe de neurones, tous disposés de la même manière ou presque dans l'espace. Par parthénogénèse, on obtient aisément plusieurs individus génétiquement identiques, appelés « isogéniques ». Il suffit alors de les couper en tranches fines, de les examiner au microscope électronique et de comparer chez chacun de ces individus l'arborisation axonale complète d'un même neurone au microscope électronique. On constate alors que, si les grandes lignes de la connectivité sont conservées, dans le détail, au niveau de la distribution des contacts synaptiques, une *variance* apparaît.

Deuxième « preuve » : l'étude de l'évolution de la connectivité cérébrale en fonction de l'expérience. On soumet un jeune chat, ou un singe nouveau-né, pendant une période sensible qui suit la naissance, à des environnements visuels artificiels différents des environnements dans lesquels ils se développent naturellement. Résultat : chez l'adulte, la spécialisation fonctionnelle de neurones individuels du cortex visuel (spécificité d'orientation, binocularité...) est grandement perturbée, le plus souvent de manière irréversible. Chez l'homme, de telles « expériences » ont lieu spontanément lorsque le nouveau-né a, par exemple, la cataracte. L'opacité du cristallin aux stades précoces du développement entraîne un déficit visuel, voire une cécité qui persiste après opération de la cataracte (si celle-ci est effectuée *après* la période sensible) se situe donc au niveau du

1. F. Harbor et C. Levinthal, E. Macagno, *Anatomy and development of identified cells in isogenic organisms*, Cold Spring Symp., Quant. Biol., 40, 1976, 321-333.

cortex cérébral. Ces expériences, parmi beaucoup d'autres, suggèrent que *l'activité* du système nerveux, au cours de son développement, contrôle la « mise au point » de la connectivité adulte. Philippe Courrèges, Antoine Danchin et moi-même ¹ avons proposé un modèle formel de l'évolution de la connectivité d'un système de neurones en développement suivant un schéma darwinien suivant ce modèle, la structure du matériel génétique ne change pas, d'où sa qualification d'« épigénèse » par stabilisation sélective de synapses. L'idée est que les déterminants génétiques engagés dans la reconnaissance entre neurones individuels faisant partie de deux groupes de partenaires cellulaires sont les mêmes ou sont très voisins. Il suffit donc de peu de gènes pour coder cette propriété de reconnaissance. A un certain stade critique, ou *période sensible*, du développement, ces deux ensembles de neurones entrent en contact. Il ne s'agit pas de contacts définis entre un neurone x du premier groupe avec le neurone y du second, mais d'une « prise de contact » exubérante, diffuse, multiple et chevauchante. Une diversité considérable de connexions se manifeste à ce stade. Le générateur de « variations darwiniennes » entre en action ! Un affinage s'ensuit, qui permet de mettre au point, par stabilisation de certaines connexions et élimination des autres, la connectivité adulte. Ce modèle permet de simuler des situations d'apprentissage simples, mais également les situations plus complexes qui se présentent au cours du développement, en particulier chez l'homme. Chez celui-ci, des vagues successives de formation et de sélection de synapses se succèdent, se déposent, s'emboîtent les unes avec les autres, par flux et reflux successifs... longtemps encore après la naissance. Il faut, bien sûr, préciser les contraintes bio-

1. J.-P. Changeux, P. Courrèges et A. Danchin, *A theory of the epigenesis of neuronal networks by selective stabilisation of synapses*, Proc. Nat. Acad. Sc. USA n° 70, 1973, p. 2974-2978.

logiques qui conduisent à la sélection d'une connexion de préférence à une autre. Ces règles de sélection doivent prendre en compte l'organisme, dans son ensemble, en interaction avec le monde extérieur.

AC : Mais pourquoi toutes ces connexions, puisqu'elles existent, ne sont-elles pas utilisées? Elles devraient l'être. Qu'est-ce qui explique la sélection?

JPC : L'activité qui circule dans le système contribue au modelage de l'état final du réseau. Et, précisément, elle n'est pas *exactement* identique pour chaque neurone. Dans le modèle proposé, des règles *locales* d'évolution déterminent l'évolution d'une synapse donnée en fonction de son état d'activité propre et de celui de la cellule sur laquelle elle se projette. Par exemple, comme je l'ai déjà mentionné, la coïncidence d'activité entre les deux cellules en contact peut entraîner la stabilisation de ce contact. L'apprentissage conduit à une relation entrée/sortie nouvelle. Après apprentissage, la même entrée produit toujours la même sortie, alors qu'avant apprentissage, plusieurs sorties sont possibles, d'une expérience à l'autre.

Ce modèle formel possède une propriété mathématique intéressante qui peut être formulée sous le nom de *théorème de la variabilité*. Il montre qu'on peut obtenir la même relation entrée/sortie après apprentissage, même si la sélection a retenu des connectivités différentes. Ce qui corrobore les observations de variance de la connectivité dont j'ai parlé tout à l'heure. On sait également que chez une majorité de personnes, les centres du langage sont distribués sur l'hémisphère gauche, chez d'autres sur l'hémisphère droit, chez d'autres encore, sur les deux hémisphères. Mais personne n'arrive à les distinguer en fonction du langage que les uns et les autres emploient. Le phénotype neural, en dépit d'une similitude remarquable de fonctions, varie donc considérablement. Nous arrivons ainsi à une conclusion extrêmement utile pour notre débat. Les mathématiciens, malgré

des différences de détail importantes de leur organisation cérébrale, arrivent à capturer par leur cerveau des objets mathématiques identiques.

Après avoir évoqué le darwinisme neural de l'évolution de la connectivité, abordons maintenant une autre évolution, de niveau plus élevé, et que l'on peut qualifier de *darwinisme mental*¹, ou darwinisme psychologique². Le darwinisme neural se manifeste principalement au cours du développement de la petite enfance ou de la vie embryonnaire. L'embryon est actif, il présente une activité spontanée qui peut intervenir dans la sélection « interne » des synapses qui assurent la coordination entre les divers centres nerveux. Le darwinisme mental, quant à lui, concerne principalement le cerveau adulte, tant au niveau de l'entendement qu'à celui de la raison. A l'échelle du temps psychologique, il produit des changements d'*efficacité synaptique*, plutôt qu'une évolution du *nombre* de connexions. Les unités de sélections ne sont plus simplement des connexions, des circuits élémentaires, mais des assemblées de neurones susceptibles d'entrer en activité de manière coordonnée. Elles se recrutent parmi les éléments déjà sélectionnés au cours du développement par darwinisme neural. Le générateur de diversité ne résulte plus de la variabilité des connexions au cours du développement, mais de l'entrée en activité, spontanée et transitoire, d'assemblées de neurones, de ce que nous avons appelé « pré-représentations ». Une activité combinatoire se développe alors, qui *anticipe* l'interaction avec le monde extérieur. Soit il y a « congruence », « résonance » entre l'état interne du système et l'état externe, et la pré-représentation est stabilisée, stockée dans le réseau. Soit il n'y a pas résonance, et aucune mise en mémoire n'a

1. J.-P. Changeux et S. Dehaene, *Neuronal models of cognitive fonctions*, Cognition, 1989 (sous presse).

2. Cette notion est déjà présente dans l'œuvre de S. Freud : voir F. Sulloway, Freud, *Biologiste de l'Esprit*, 1981, Fayard, Paris, p. 244.

lieu. Cette mémorisation change les efficacités synaptiques qui vont stocker une certaine configuration du système.

L'enchaînement des représentations mentales évoquées, par la raison, dans le « compartiment de travail » de la mémoire à court terme procède d'un darwinisme mental similaire. Ce modèle s'applique-t-il au travail du mathématicien ? On peut imaginer qu'au cours de la période d'« incubation », des représentations variées des objets mathématiques se succèdent de manière transitoire, s'enchaînant un peu au hasard. Puis, une sorte de sélection *interne* par résonance a lieu entre représentations ou pré-représentations. Elle aboutit à un « objet-résultat » qui concorde avec le problème posé, avec « l'intention » à laquelle la réponse doit être apportée. Au stade où nous en sommes, je ne pense pas que ces idées encore très schématiques puissent être présentées sous une forme plus précise.

Stanislas Dehaene, Jean-Pierre Nadal et moi-même¹ avons élaboré un modèle encore très élémentaire de réseau de neurones disposés en couches successives qui est capable de reconnaître des séquences, des mélodies, de les mettre en mémoire et, éventuellement, de les produire. Ce modèle rend assez bien compte de l'apprentissage du chant chez certains oiseaux (*voir figure 24*). Nous en sommes à peu près là, mais je pense qu'un jour, il sera possible de modéliser dans le détail certaines étapes du déroulement de la pensée.

7. *Darwinisme mental et création mathématique*

JPC : Je te propose donc de considérer que certaines formes de l'activité mentale du mathématicien, comme de

1. S. Dehaene, J.-P. Changeux et J.-P. Nadal, *Neural networks that learn temporal sequences by selection*, Proc. Nat. Acad. Sc. USA n° 84, 1987, p. 2727-2731.

la pensée en général, suivent une sorte d'évolution darwinienne...

AC : On pourrait également former l'hypothèse de la dualité entre le processus aléatoire du darwinisme et ma croyance dans l'existence indépendante d'une réalité mathématique brute. Sa cohérence et son harmonie sont l'antidote de l'aléatoire. Plusieurs raisonnements relativement aléatoires aboutissant au même résultat indiquent qu'on est en bonne voie. Au troisième niveau, la cohérence inexplicable de la réalité mathématique permet précisément, me semble-t-il, à plusieurs assemblées indépendantes de neurones de n'entrer en résonance que lorsqu'elles sont en harmonie.

JPC : Oui, c'est la combinatoire des pré-représentations.

AC : Il faut avoir postulé qu'il existe, indépendamment du cerveau, un monde dont la cohérence peut être capturée, peut-être, à travers la résonance des mécanismes aléatoires.

JPC : J'allais te suggérer une idée de ce style. Allons plus loin dans la définition des objets mathématiques comme objets mentaux, et considérons-les d'abord comme des représentations mentales de caractère privé, comme des états physiques, observables avec une caméra à positons.

AC : En elle-même, une représentation mentale ne signifie rien...

JPC : Elle prend un « sens » explicite dès qu'elle est communiquée. Les objets mathématiques sont en effet des représentations mentales, qui possèdent la propriété essentielle d'être communicables d'un individu à l'autre, à la différence des états « ineffables » des grands mystiques ou des aliénés. Ils deviennent des représentations publiques. Les objets mathématiques sont transmissibles, de manière « rigoureusement exacte », d'un cerveau à l'autre, et manipulables, de manière comparable, par des individus génétiquement et épigénétiquement distincts.

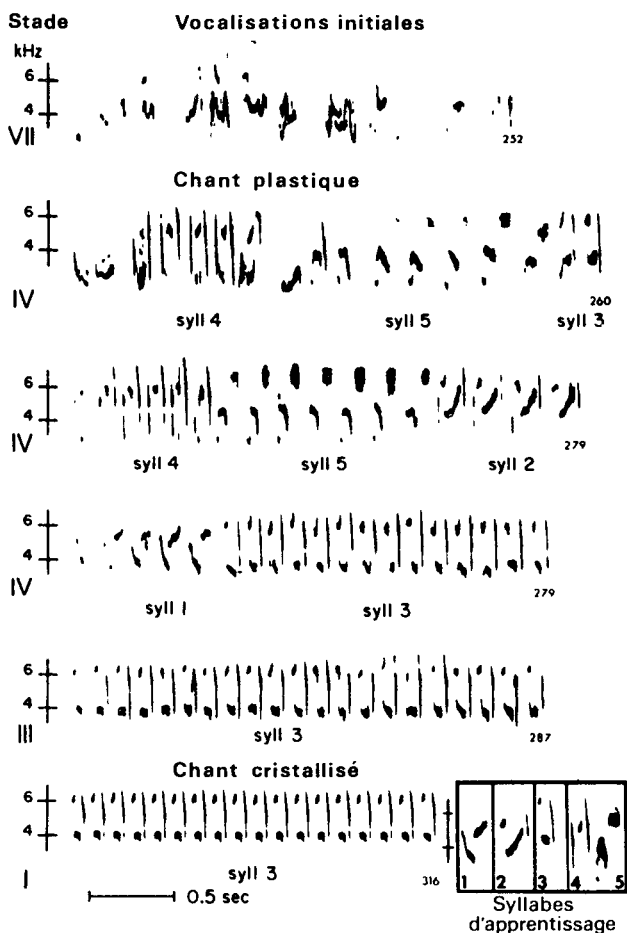


FIGURE 24

L'apprentissage du chant chez le moineau des marais. Les tracés représentent la fréquence du son produit en fonction du temps. Les syllabes d'apprentissage sont représentées en bas à droite. Le jeune oiseau les entend et les mémorise entre les jours 22 et 62 après l'éclosion. Près de 200 jours plus tard, l'oisillon produit ses premières vocalisations. Elles se regroupent en syllabes qui reproduisent les syllabes d'apprentissage entendues 7 mois plus tôt. Seule la syllabe n°3 persiste dans le chant de l'adulte. La cristallisation du chant s'accompagne d'une perte de syllabes (ou attrition syllabique) qui témoigne du caractère « sélectionniste » de l'apprentissage. (D'après

P. Marler & S. Peters, 1982, Subsong and plastic song : their role in the vocal learning process. D.E. Kroodsma & E.H. Miller ed. « Acoustic communication in birds », vol 2, 25-50)

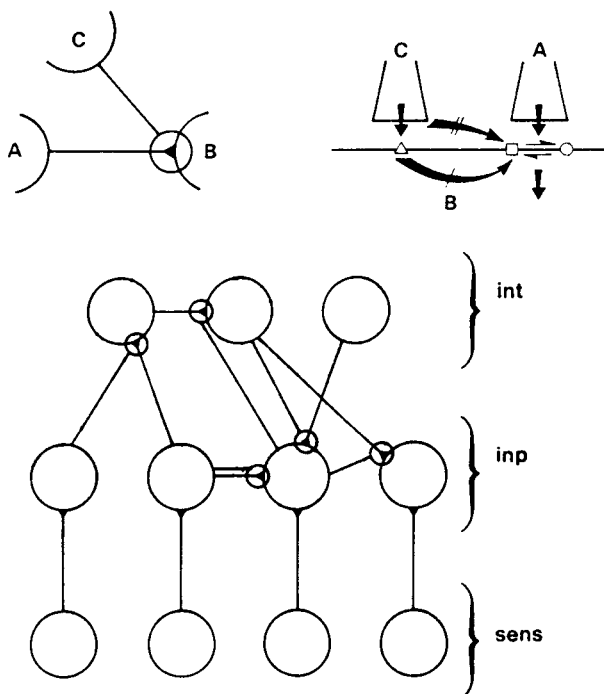


FIGURE 24bis

Réseau de neurones formels susceptibles de reconnaître, produire et mettre en mémoire par sélection des séquences temporelles de « représentations ». L'architecture est très simple : 3 couches de neurones (sensoriels (sens)), d'entrée (inp), et interne (int)) subdivisés en groupes de neurones auto-excitateurs (cercles) codant pour les « représentations » ; enfin les neurones sont réunis entre eux par des synapses en triades ABC dont les efficacités sont modulables chimiquement. Pour plus de détails, voir Dehaene et al. (1987).

Certains anthropologues, comme Sperber ¹, distinguent plusieurs types de représentations publiques. Des représentations de « premier ordre », qui expriment, par exemple, que le pain est comestible, que le lion est dangereux, que les plantes sont vertes. Elles sont enregistrées dans la mémoire à long terme et excluent toute inconsistance empirique ou toute contradiction entre elles. Bien que factuelles, elles ont donc valeur d'universalité, puisqu'elles ont été vérifiées par tout le monde. Pour Sperber, les représentations de second ordre sont des « représentations de représentations », des relations entre faits et états mentaux ou entre états mentaux intersubjectifs. Il les sépare en deux : les croyances et les modèles scientifiques. J'en ajouterai une troisième, celle des représentations artistiques ².

Les croyances sont, par définition, variables. Elles sont pourtant transmises de manière *autoritaire* comme des vérités, ce qui est, comme le dit Sperber, une provocation permanente contre le sens commun. A côté des croyances, se développent les hypothèses, les modèles scientifiques ou bien les objets mathématiques, qui se définissent par leur caractère cohérent, non ambigu, non contradictoire, prédictif et génératif. En plus de leur adéquation au réel, parce que falsifiables et éventuellement révisables, ils s'opposent aux croyances, intouchables dans un corpus théologique donné ! Celles-ci suivent bien une évolution qui peut s'interpréter en termes darwiniens, mais elle se distingue de celle des objets mathématiques. Voilà donc définis les traits propres aux objets mathématiques en tant que représentations publiques de second ordre, en tant que représen-

1. D. Sperber, *Anthropology and psychology : towards an epidemiology of representations*, Man (N.S.) n° 20, 1984, p. 73-89.

2. J.-P. Changeux, *Raison et plaisir*, Catalogue de l'Exposition : De Nicolo del Abate à Nicolas Poussin : Aux sources du Classicisme, Commissaires J.-P. Changeux et B. Grinbaum, nombreux auteurs, 20-32, Musée de Meaux.

tations scientifiques aussi décantées, aussi « pures » que possible.

La composante émotionnelle paraît bien plus importante que la composante rationnelle dans la sélection et la propagation des croyances. Qu'en est-il de la sélection des objets mathématiques? Parmi les étapes du travail du mathématicien, tu as cité l'illumination qui se produit après une phase d'incubation lors de laquelle la combinatoire darwinienne semble avoir lieu. On peut considérer que l'illumination coïncide avec une entrée en résonance de représentations mentales entre elles. Or, le cortex frontal, où cette résonance a vraisemblablement lieu est très directement lié au système limbique, engagé, quant à lui, dans les états émotionnels (*voir figure 25*). Notre cortex frontal non seulement élabore des stratégies cognitives, mais est aussi capable, par le truchement de connexions très riches entre cortex frontal et système limbique (*voir figure 26*), de développer des stratégies émotionnelles. Je pense que le mathématicien doit lui aussi développer, en même temps que ses stratégies rationnelles, des stratégies émotionnelles qui lui font espérer arriver au résultat. Lors de l'illumination, les résonances débordent du cortex frontal jusqu'au système limbique, de sorte qu'on pourrait même aller jusqu'à dire que l'état émotionnel contribue à l'évaluation...

AC : Tout à fait. Et c'est très important.

JPC : Cette *fonction d'évaluation*, susceptible de reconnaître une « harmonie » entre le sujet et son environnement, ou « une harmonie » intérieure entre plusieurs représentations, peut s'interpréter comme un système de plaisir ou comme un système d'alarme.

Enfin, il faudrait distinguer les conditions dans lesquelles l'illumination se produit et celles de sa transmission d'un mathématicien à l'autre. Il s'agit de processus différents. La création se distingue du transfert de connaissances.

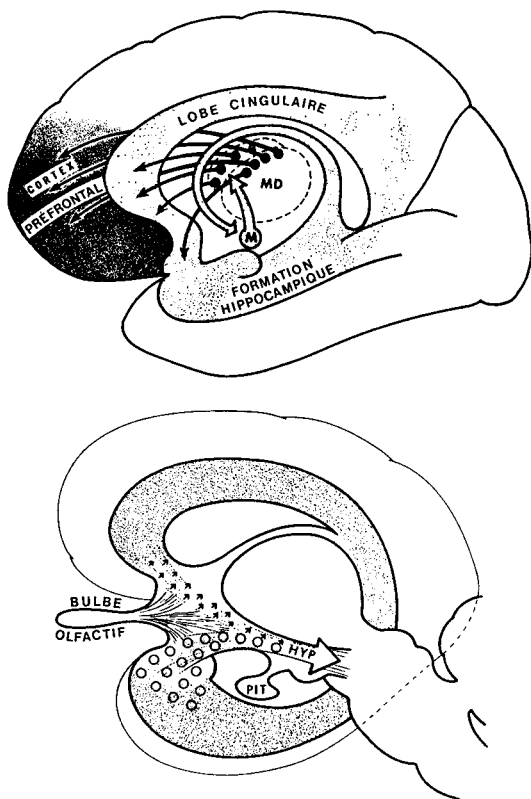


FIGURE 25

Système limbique et plaisir. Représentation très schématique du système limbique et de son organisation en circuit, décrit pour la première fois par Papez. Il se compose, en particulier, de l'hippocampe qui reçoit des informations du néocortex, de l'hypothalamus (Hyp), dont les corps mamillaires (M) font partie, des noyaux antérieurs (A) et postérieur du thalamus (MD). Ceux-ci se projettent respectivement sur le cortex pré-frontal et le lobe cingulaire, dont la forme circulaire ressemble à un « limbe », d'où le nom de « grand lobe limbique » donné par Broca à celui-ci.

La stimulation électrique de points précis du système limbique entraîne une conduite d'auto-stimulation et donc une sensation de plaisir. Les flèches obliques indiquent une réponse manifestée par une érection du pénis chez le mâle. (D'après P. Mac Lean, 1973, *A Triune Concept of the Brain and Behavior*, University of Toronto Press.)

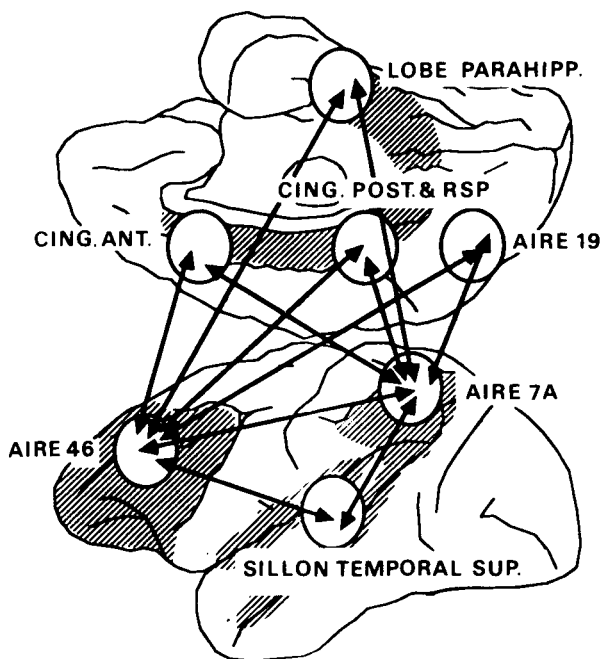


FIGURE 26

Réseau de connexions anatomiques établies chez le singe entre le lobe frontal (aire 46), le lobe temporal (sillon temporal supérieur) et le lobe pariétal (aire 7A) avec le système limbique (lobe cingulaire antérieur (cing. post.), lobe cingulaire postérieur (cing. post.), et lobe parahippocampique) chez le singe. Le bas de la figure représente la face externe de l'hémisphère gauche, le haut de celle-ci sa face interne. Le système limbique est présent essentiellement sur la face interne des hémisphères cérébraux, dans la partie médiane du cerveau. Les connexions réciproques néocortex/système limbique assurent le lien entre cognition et émotion. (D'après Goldman-Rakic P., 1988, Topography of Cognition : Parallel distributed networks in primate association cortex, *Ann. Rev. Neurosc.*, 11, 137-156.)

Néanmoins, pour que la communication se fasse, le cerveau du receveur doit posséder une compétence particulière...

AC : Bien sûr.

JPC : Un certain niveau de compétence est requis pour que le receveur accepte, ou rejette, l'objet mathématique ou la démonstration qui lui sont proposés. Il faut donc prendre en compte cette compétence, représentative du corpus mathématique existant. L'adoption de la proposition nouvelle par la communauté des mathématiciens signifie, en particulier, la cohérence, l'intégration de celle-ci à ce corpus. La cohérence interne des objets mathématiques qui t'étonne tant s'élabore donc progressivement.

AC : Nous en construisons sans doute progressivement une copie dans notre cerveau par un processus d'imagerie mentale. Mais cela ne remet pas en cause l'existence de la réalité mathématique elle-même.

JPC : Elle se construit progressivement par ces ouvertures aux variabilités et par entrées en résonance avec le reste de ce corpus. C'est pourquoi je nie une réalité des mathématiques antérieure à l'expérience qu'on en a. Leur cohérence, au lieu d'être *a priori*, me paraît *a posteriori* et résulter tout simplement de leur *non contradiction*. C'est également pour cela que Morris Kline intitule son ouvrage sur l'histoire récente des mathématiques : *La fin de la certitude*!

AC : Nous en revenons à notre débat de départ. Je pense qu'il est temps de le dépasser.

V

Darwin chez les mathématiciens

1. L'utilité du schéma darwinien

JEAN-PIERRE CHANGEUX : Le darwinisme en mathématiques me paraît une idée neuve. Avant de la développer, il serait utile de préciser à nouveau les niveaux dont nous avons parlé l'un et l'autre, pour que nous puissions définir plus concrètement les « évolutions » susceptibles, pour chacun, d'intervenir.

ALAIN CONNES : Je n'ai pas grand-chose à ajouter. De façon assez formelle, on peut distinguer, pour les besoins de la discussion, trois niveaux. Mais je ne prétends nullement qu'ils aient un sens absolu. Le premier, tout d'abord, se définit par la faculté de calculer, d'appliquer une recette donnée, à la fois vite et bien. Il est déjà présent dans les ordinateurs actuels.

JPC : C'est le niveau des opérations symboliques.

AC : Oui, mais ces opérations peuvent être très compliquées. Quel que soit le degré de complexité, pourtant, la

recette est toujours donnée à l'avance. Elle n'est nullement comprise. Aucune variation, aucun changement de stratégie ne sont donc possibles.

JPC : C'est le niveau symbolique, celui de l'entendement, qui se situe, selon Kant, entre la sensibilité et la raison.

AC : Au deuxième niveau au contraire, il est possible, pour une fin déterminée, comme, par exemple, la résolution d'un problème, de choisir une stratégie, et d'en changer suivant le résultat. Lorsqu'une erreur se produit, il est possible de faire des comparaisons avec d'autres calculs. Ce niveau suppose donc la compréhension du mécanisme utilisé. Par exemple, pour une division, on comprend pourquoi on effectue telle ou telle opération plutôt que telle autre, ou, pour prendre une image un peu exagérée, lorsqu'on fait une retenue dans l'addition, on comprend qu'on utilise un 2 cocycle de groupe. Il faut donc avoir formalisé les opérations utilisées, les avoir hiérarchisées en fonction du but auquel adapter la stratégie choisie, et, pour ce faire, adhérer vraiment à ce que l'on fait. En mathématiques, c'est ce qui permet souvent de résoudre un problème, lorsqu'il n'est pas trop difficile ou ne nécessite pas d'idées nouvelles. A condition bien entendu qu'il ne relève pas du premier niveau, c'est-à-dire ne soit pas un simple calcul, ou la simple application d'une recette.

JPC : C'est une forme, peut-être inférieure, de la raison, qui pourrait se comprendre comme la raison tactique de Granger. Elle fait intervenir la mise en application d'une stratégie, et, éventuellement, la recherche d'une nouvelle tactique si la première échoue.

AC : Le fil de la pensée, au deuxième niveau, n'est jamais perdu. C'est ce qui me paraît le distinguer très bien du troisième. A aucun moment, il n'existe de distance entre le fonctionnement du cerveau et l'objet auquel il s'applique.

JPC : C'est la définition que Granger donne de la raison

tactique, me semble-t-il. On change de tactique, on change de moyen, de méthode, mais on garde la même intention mathématique. Au contraire, le troisième niveau permet de changer complètement de stratégie, de fin poursuivie.

AC : Attention. Ce n'est pas la distinction que je voulais faire. Pour moi, on peut définir le troisième niveau de la manière suivante : l'« esprit », ou la « pensée », est occupé à une tâche différente, tandis que, de manière interne, subconsciente pourrait-on dire, le problème est en train de se résoudre. L'essentiel est justement cette dissociation entre la pensée présente, active, et un fonctionnement du cerveau qui n'est pas apparent...

JPC : Je ne pense pas qu'il faille considérer que le fait qu'une opération soit consciente relève d'un niveau particulier. Il s'agit plutôt d'un mode de perception « intérieur » de ce qui se produit. Ce qui caractérise le troisième niveau semble être, au contraire, l'accès possible à une illumination qui permet un changement global de stratégie. Se crée alors un nouveau cadre de pensée dans lequel, ensuite, pourra s'appliquer une tactique nouvelle. Nous sommes finalement d'accord sur ces trois niveaux. Dans ces conditions, le schéma darwinien te paraît-il utile ?

AC : Pour apprécier son efficacité, il me manque la définition plus ou moins précise d'une fonction d'évaluation. Celle-ci permettrait, par exemple au deuxième niveau, d'avoir l'intuition qu'une stratégie est meilleure qu'une autre et de faire des choix. Pour que le darwinisme soit efficace, il faut tout de même que le cerveau arrive à sélectionner parmi les diverses possibilités, ou parmi les divers agrégats de neurones, ceux qui fonctionnent le mieux. Il faut qu'il puisse choisir en fonction du but proposé. Mais comment concevoir l'idée, même imprécise, même vague, d'une fonction d'évaluation un peu analogue, et peut-être adaptable dans certains ordinateurs, pour choisir une stratégie meilleure qu'une autre ?

JPC : La balle est dans ton camp.

AC : Si l'on prend pour principe celui que j'ai adopté dès le début, à savoir l'existence indépendante de la réalité mathématique, on peut proposer un certain nombre d'idées, à titre au moins d'exemples qui doivent être confrontés avec l'expérience, avec la réalité. On peut ainsi concevoir que la cohérence interne des mathématiques est un guide, au sens où toute structure s'oppose à l'aléatoire. Ainsi, il n'est pas impossible que lorsque le cerveau essaie de construire une représentation imagée d'une réalité mathématique, ce soit sa cohérence qui joue le rôle de mécanisme de sélection.

JPC : Cela n'implique en aucune manière que les mathématiques préexistent. Tu en parles comme d'un guide de pensée. La fonction d'évaluation est une fonction de vérification de l'intégration dans une structure cohérente, non contradictoire. Cette vérification a lieu dans notre cerveau, où se trouvent stockées, dans la mémoire à long terme, un certain nombre de représentations mathématiques. Celles-ci entrent en quelque sorte en résonance, quand ce nouvel objet apparaît. Une mise en activité globale se produit alors.

AC : Je pensais à une cohérence interne.

JPC : Oui. Je veux dire qu'elle est interne à la fois au cerveau et aux mathématiques, parce que les mathématiques sont *dans* le cerveau du mathématicien, en particulier dans sa mémoire à long terme. Elles sont représentées par des indices dont l'ensemble est soudain réuni par un nouvel objet mathématique. Tout ce monde entre alors en activité de manière coordonnée. Les éléments du puzzle étaient presque tous présents, mais une pièce manquait pour que l'ensemble s'organise. La figure apparaît subitement lorsqu'on ajoute la nouvelle pièce.

AC : Mais revenons à l'opposition entre, d'un côté, le désordre et, de l'autre, l'organisation. La réalité mathématique, de par sa structure, son harmonie interne, est une réserve inépuisable d'organisation. En choisissant au hasard

des formules, on n'obtiendra de résonance entre elles que si elles ont, ensemble, une certaine cohérence. Les mathématiques ont justement pour fonction de montrer l'existence de cette cohérence. On pourrait penser à divers groupements de neurones en activité, qui n'arrivent à entrer en résonance que lorsque se produit ce genre de phénomène de cohérence. Cette idée est pour l'instant vague, mais elle mérite d'être précisée.

JPC : L'accès à cette cohérence fait intervenir, pendant la période d'incubation, une variabilité. A cet égard, le cerveau ne fonctionne pas comme un ordinateur ou comme une machine à jouer aux échecs. Toutes les possibilités ne sont pas prises en considération et évaluées. Au contraire, s'il y a combinatoire, celle-ci semble porter sur un relativement petit nombre d'objets de pensée.

AC : Une fois établie dans le cerveau une structure minimale, un modèle même très primitif donnant une image mentale d'une réalité mathématique même simple, il n'est pas très difficile d'imaginer, à l'intérieur du cerveau, un mécanisme d'évolution des circuits qui permette de créer des structures plus élaborées. Prenons comme exemple le raisonnement par analogie. Ce type de raisonnement revient, à partir d'une structure syntaxique simple, à créer un modèle semblable dont les éléments ont une interprétation sémantique différente. Si l'on teste la compatibilité de cette nouvelle structure avec la réalité mathématique, on peut la modifier et la rendre plus performante. La résolution d'un problème ne résulte donc pas nécessairement d'une suite d'essais aléatoires. On peut très bien avoir directement accès à un nombre plus réduit de possibilités de solution, grâce à une analogie construite à partir d'un modèle passé. Nous parlions des échecs. Je crois que les grands joueurs d'échecs arrivent, grâce à cette intuition, à restreindre le nombre de coups qu'ils ont à considérer, alors que l'ordinateur, lui, en essaie des millions.

JPC : Les psychologues ont étudié le cas des grands maîtres d'échecs et analysé leurs stratégies ¹. Il semble que ceux-ci apprennent en quelque sorte un langage nouveau qui contient une série importante de possibilités, de coups à jouer, correspondant chacun à un mot. Leur nombre serait d'environ sept à dix mille, comme pour un dictionnaire de français ou d'anglais. Au lieu d'analyser d'une manière systématique et combinatoire la distribution des pions sur l'échiquier, le grand maître fait appel à sa mémoire pour élaborer la stratégie adéquate. Donc, plutôt que d'inventer en permanence de nouvelles stratégies, il raisonne de préférence à partir des images et des stratégies qu'il a mémorisées.

AC : La notion de stabilité des configurations, des formes, me paraît ici très importante. Le cerveau perçoit de manière semblable certaines formes qui, codées strictement, sont différentes. Par exemple, au jeu d'échecs, le grand maître parvient à découvrir et à classer, grâce à ce mécanisme, un petit nombre d'« attracteurs » parmi un grand nombre de configurations qui, éloignées par leur position, sont voisines dans son esprit. Ce mécanisme spécifiquement mental, auquel pour lors n'ont pas accès les ordinateurs, lui permet donc de limiter son problème à un petit nombre de possibilités de solution. L'intelligence artificielle cherche d'ailleurs aujourd'hui à imiter ce processus grâce à la topologie dynamique.

2. *Coder des formes stables*

JPC : La mémoire à long terme est donc hiérarchisée. Elle n'a rien d'un dictionnaire classé par ordre alphabétique. Au contraire...

1. A. de Groot, *Thought and Choice in Chess*, Basic Books, New York, 1965. Voir également D. Hofstadter, Gödel, Escher, Bach, Interéditions, Paris, 1985.

AC : Elle est hiérarchisée, me semble-t-il, par des mécanismes de topologie.

JPC : L'organisation de la mémoire à long terme est un problème théorique fondamental pour les neurobiologistes. On parle d'arbres sémantiques, de classifications hiérarchiques...

AC : Avant de me lancer dans quelques explications sur la topologie, je ne peux m'empêcher de commencer par une mise en garde. Un excellent mathématicien de mes collègues décida un jour de s'occuper de psychanalyse : il pensait peut-être que la topologie pouvait être un outil intéressant pour les investigations des psychanalystes. Il racontait qu'un jour Lacan, qui venait d'apprendre ce qu'est un espace compact, avait expliqué dans son cours que Don Juan était compact, en montrant qu'on peut extraire un recouvrement fini de tout recouvrement ouvert... Certains membres du groupe de Lacan utilisèrent ainsi le langage mathématique, sans en connaître le contenu, pour exercer un ascendant psychologique sur d'autres collègues tout aussi ignorants des mathématiques. Il est clair que le monde chimérique ainsi créé ne correspond à aucune réalité. Notre discussion doit absolument éviter de donner lieu à ce genre de fausses interprétations. En particulier, je ne prétends avoir aucune sorte de compréhension nouvelle du fonctionnement cérébral. Je pense seulement qu'il serait bon que certaines notions élémentaires de topologie, que je vais m'efforcer d'expliquer en détail, soient mieux connues des scientifiques neurobiologistes comme toi. Mais pourquoi la topologie ? Comme tu l'as expliqué, la conformation du cerveau n'est pas identique d'un cerveau à l'autre, pas plus que la perception d'un objet extérieur. Mais les propriétés sur lesquelles on s'accorde ont un caractère d'invariance, de « stabilité structurelle » pour citer Thom, dont un cadre théorique comme la topologie rend assez bien compte.

JPC : Les représentations mentales, les objets de

mémoire, sont codés dans le cerveau comme des formes, dans le sens de la *Gestalt-Theorie*, en dépit de l'importante variabilité des synapses qui les stockent. Il existe donc des processus d'implémentation des invariants perceptuels dans le système nerveux. C'est un premier problème. L'autre concerne la manière dont ces représentations sont classées dans la mémoire. Il faut les séparer. Commençons par le premier...

AC : Je vais tout d'abord tenter d'expliquer la notion fondamentale de topologie simpliciale, dont le cas le plus simple est la notion d'arbre que tu as mentionnée. Elle a pour but l'étude des invariants topologiques d'objets appelés complexes simpliciaux. Un complexe simplicial est un ensemble fini de points que j'appellerai « sommets ». Tu peux te les représenter comme des neurones pris en un agrégat assez complexe. La structure de cet objet est donnée par un sous-ensemble, que j'appellerai « Delta 1 », de l'ensemble des couples de sommets. Nous appellerons ses éléments les « arêtes », et pour continuer la comparaison avec les neurones, tu peux te représenter une arête comme une connexion entre deux neurones. Mais, sauf si le complexe simplicial est unidimensionnel, la structure ne s'arrête pas là. En général, on doit donner pour tout entier n plus petit que la dimension un sous-ensemble « Delta n » de l'ensemble des n -uples de sommets. Par exemple, si le complexe simplicial est de dimension 2, il faudra donner non seulement les arêtes, mais aussi les triangles. La seule règle de compatibilité est la nécessité pour les côtés qui bordent un triangle d'être des arêtes. Cela signifie qu'un triangle ABC n'appartient au complexe $ABC \in \Delta^2$ que si ses trois arêtes appartiennent au complexe : $AB \in \Delta^1$, $BC \in \Delta^1$, $AC \in \Delta^1$. Mais la réciproque n'est pas vraie. De même que si A et B sont deux sommets, l'arête qui les joint n'appartient pas nécessairement au complexe. A partir de cette donnée, on peut

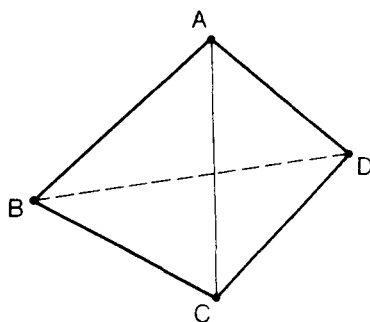
mettre en œuvre le potentiel considérable de la topologie simpliciale.

Les complexes simpliciaux de dimension 1 ne sont pas très intéressants. Ainsi, le groupe fondamental de l'espace topologique associé est toujours un groupe libre. Je vais prendre quelques exemples de complexes simpliciaux de dimension plus grande, sans pour autant chercher, du moins pour l'instant, à leur donner une signification ou une réalisation quelconque. Mais, pour que tu puisses te les représenter facilement, j'utiliserai le terme de neurones pour les sommets de mon complexe simplicial, celui de connexions simples pour les arêtes qui relient un neurone à un autre, et celui de connexions multiples pour parler d'un n -uple de neurones. Mon premier exemple sera un complexe simplicial dont la topologie est celle de la sphère à deux dimensions. Il est formé de quatre neurones A, B, C, D. Tous les couples (AB, AC, BD...) sont reliés par des arêtes. Tous les triplés sont reliés par un triangle. Par contre, la dimension est égale à 2, et il n'existe aucune liaison au-delà des liaisons triangulaires. Je vais maintenant décrire un autre complexe simplicial équivalent au précédent, c'est-à-dire définissant le même objet topologique, mais dont le nombre de sommets sera différent. On ajoute un sommet E. Les arêtes sont les anciennes, plus celles qui joignent E aux sommets A, B et C, pas celle qui joint E à D. Les triangles sont tous les triangles rendus possibles par l'existence de leurs arêtes (*voir figure 27*), sauf le triangle ABC. Par exemple AEB est un triangle, mais AED n'est pas un triangle, car ED n'est pas une arête. Le complexe simplicial obtenu est de dimension 2, et l'espace topologique associé est homéomorphe à l'espace topologique associé au premier complexe simplicial. Ces deux espaces sont homéomorphes à la sphère à deux dimensions. Le passage du premier complexe simplicial au second est ce qu'on appelle une « subdivision barycentrique ».

$$\triangle^1 = (A, B, C, D)$$

$$\triangle^2 = (AB, AC, AD, BC, CD)$$

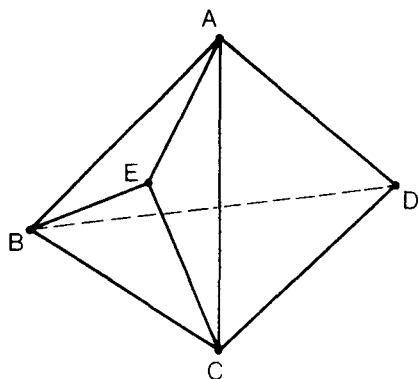
$$\triangle^3 = (ABC, ABD, ACD, BCD)$$



$$\triangle^1 = (A, B, C, D, E)$$

$$\triangle^2 = (AB, AC, AD, BC, BD, \\ CD, AE, BE, CE)$$

$$\triangle^3 = (ABE, ACE, BEC, ABD, \\ ACD, BCD)$$



COMPLEXE SIMPLICIAL

RÉALISATION GÉOMÉTRIQUE

FIGURE 27

Dans la comparaison avec les assemblées de neurones, la première difficulté pour savoir s'il existe des assemblées de neurones de dimension plus grande que 1 consiste à détecter, à définir correctement, les liaisons triples ou triangles dans le complexe simplicial. Cela ne peut se faire qu'au prix d'un dialogue avec l'expérience, ou bien en construisant une machine qui, au lieu d'utiliser seulement le mécanisme de classification fondé sur les arbres, serait capable d'exploiter les ressources beaucoup plus riches de la topologie de plus grande dimension.

JPC : Je pense que c'est une idée intéressante. On sait que les neurones possèdent, individuellement, des dizaines de milliers de connexions qui, vraisemblablement, peuvent prendre part à des représentations différentes. Le schéma que tu décris permettrait d'utiliser cette possibilité... Considérons, par exemple, l'encodage par mon cerveau d'une figure particulière, disons de ton visage. Le problème des dimensions devient critique. Nous avons déjà évoqué ensemble les représentations mentales, considérées comme des états physiques définis par l'activité de populations déterminées de neurones. Mais ce point de vue n'est pas universellement partagé. Barlow ¹ a défendu la théorie alternative, suivant laquelle, dans le cerveau, chaque neurone possède une spécificité fonctionnelle extrêmement élevée qui suffit pour coder une « représentation » aussi singulière que sa propre grand-mère ou une Volkswagen jaune. C'est la théorie des « *grandmother cells* », ou cellules grand-mères. Quelques données expérimentales vont dans ce sens. Dans le cortex pariéto-temporal, chez le singe ², on peut enregistrer des neurones individuels qui servent à coder la recon-

1. H. Barlow, *Simple units and sensations: a neuron doctrine for perceptual physiology?*, in *Perception* 1, 1972, p. 371-394.

2. D.I. Perrett, A.J. Mistlin & A.J. Chitty, *Visual neurones responsive to faces*, Trends in Neurosc. 10, 1987, p. 358-364. R. Desimone, T.D. Albright, C.G. Gross & C. Bruce, *Stimulus selective perspectives of inferior temporal neurones in the macaque*, J. Neurosc., 4, 1984, p. 2051-2062.

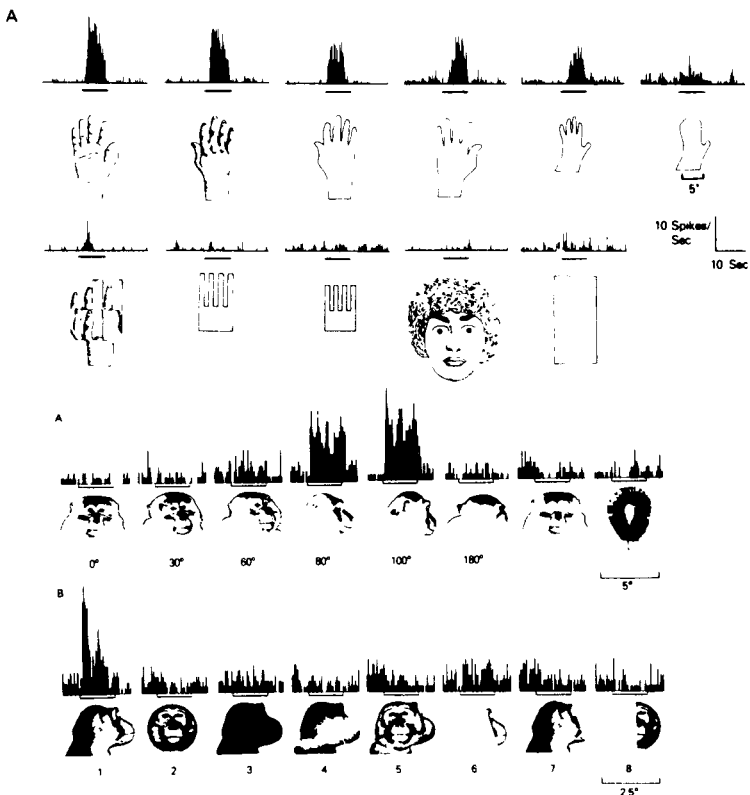
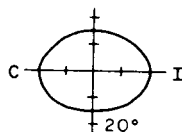


FIGURE 28

Spécificité de la réponse de neurones individuels du cortex temporal à des objets très complexes chez le macaque. La réponse de neurones individuels est enregistrée à l'aide d'une microélectrode chez le singe éveillé. Chaque influx nerveux est figuré par une barre verticale de longueur constante (ci-contre). La fréquence de ces impulsions pour un laps de temps fini est représenté par un trait de longueur variable (ci-dessus). Les spécificités des neurones enregistrées en trois points différents du cortex temporal sont respectivement : réponse à un visage de face, réponse à un visage de profil, réponse à une main. On notera que la présence des yeux est nécessaire à la réponse au visage de même que la distinction des doigts pour la réponse à la main (d'après C.G. Gross, C.J. Bruce, R. Desimone, J. Fleming, et R. Gattas, 1981, Critical visual areas of the temporal lobe, in C. Woolsey ed. *Cortical Sensory Organization*, vol. 2, 187-216, Humana Press : Clifton N.J.; R. Desimone, T. Albright, C. Gross & C. Bruce, 1984, Stimulus-selective Properties of inferior temporal neurons in the macaque, *J. Neuroscience*, 4, 2051-2062).



2 sec

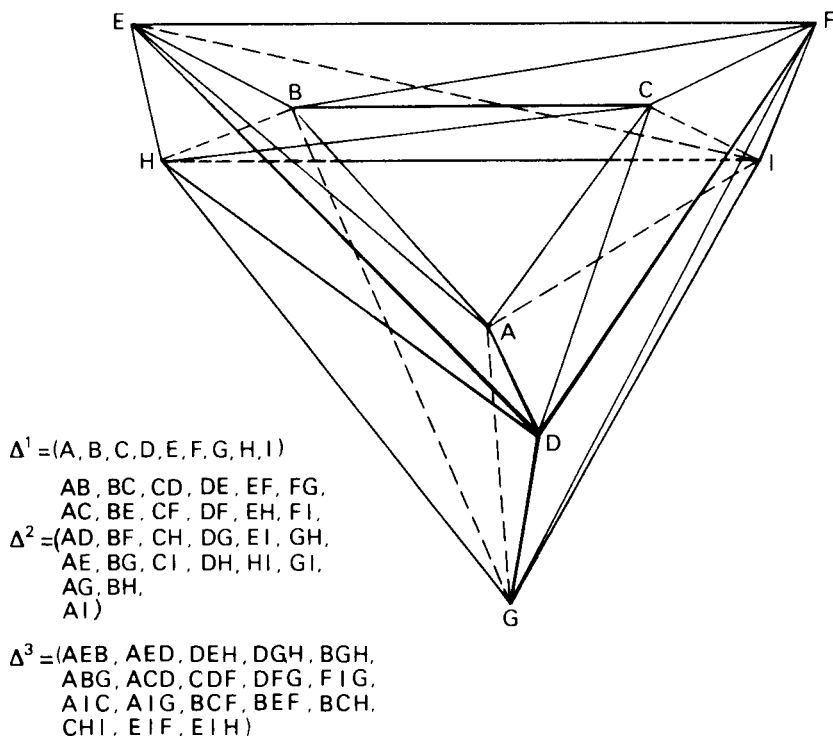


naissance des visages et même certains traits du visage (*voir figure 28*). Certains neurones répondent au visage de face *ou* de profil, d'autres au visage de face *et* de profil; d'autres répondent au visage *avec* les yeux, mais pas *sans* les yeux, au visage d'un des expérimentateurs, mais pas de l'autre; certains d'entre eux sont même sensibles à l'orientation du regard de l'expérimentateur qui les observe. Cela te donne une idée de la spécificité fonctionnelle extrêmement fine de certains neurones. Mais il ne faut pas aller trop loin, car s'il n'y avait effectivement dans le cortex temporal qu'*un seul neurone* qui code pour chacun de ces traits, on aurait très peu de chance de l'enregistrer. Le fait qu'on puisse réaliser des enregistrements reproductibles prouve qu'il existe des populations entières de neurones qui possèdent ces spécificités. Vraisemblablement, ce sont des *ensembles* de neurones hautement différenciés qui participent à la reconnaissance du visage. Et les neurones qui, individuellement répondent à des traits aussi singuliers, sont, en fait, connectés à d'autres ensembles de neurones qui se trouvent dans les aires visuelles secondaire et primaire, eux-mêmes mis en activité par les neurones de la rétine. On a donc affaire, dans le cas du système nerveux, à des systèmes à la fois hautement hiérarchisés et hautement parallèles. Je ne suis donc pas certain que ta suggestion empruntée à la topologie simpliciale puisse être mise en application dans ce cas précis.

AC : Je ne sais pas si mes remarques ont vraiment beaucoup à voir avec ce problème de la mémorisation.

JPC : Il ne s'agit pas simplement de mémorisation. Mais de la manière dont les informations sont stockées dans le système nerveux. Le problème du codage des représentations mentales est un problème de topologie.

AC : Oui, mais j'en suis venu à la topologie pour une raison assez différente. Nous parlions de l'existence d'une grande diversité, mais aussi d'une certaine invariance, dans la manière dont le cerveau est construit d'un individu à un



COMPLEXE SIMPLICIAL

RÉALISATION GÉOMÉTRIQUE

FIGURE 29

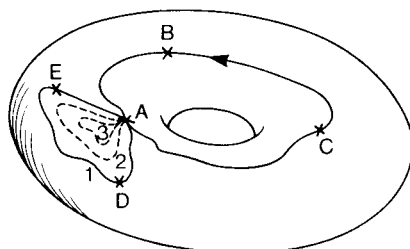


FIGURE 30

Le lacet ACB représente un élément non trivial du groupe fondamental du complexe simplicial de la figure 29. Le lacet AED représente l'élément trivial car on peut le déformer comme indiqué sur la figure dans les étapes 1, 2 et 3.

autre. La topologie est précisément le cadre idéal pour comprendre ce genre de phénomène, puisque le même objet topologique peut avoir de nombreuses réalisations différentes. Il peut provenir de nombreux complexes simpliciaux différents, tout en gardant intactes ses propriétés topologiques. La topologie simpliciale est donc un moyen idéal pour coder par exemple la notion de forme, pour autant que l'on n'insiste pas encore sur la partie quantitative de sa géométrie. Pour prendre l'exemple le plus simple d'un invariant qui ne change pas quand on remplace un complexe simplicial, disons de dimension 2, par un autre qui décrit le même objet topologique, je vais t'expliquer ce qu'est la caractéristique d'Euler et Poincaré. Il s'agit d'un nombre qui est simplement le nombre de sommets, moins le nombre d'arêtes, plus le nombre de triangles. Il n'est pas difficile de vérifier que lorsqu'on fait l'opération de subdivision barycentrique que j'ai expliquée tout à l'heure, on ajoute un sommet, trois arêtes, et deux triangles, ce qui ne change pas le nombre que je viens de définir. Si l'on calcule cette caractéristique d'Euler-Poincaré pour la sphère à deux dimensions, on a quatre sommets, quatre triangles, et six arêtes. Elle est donc égale à 2. On pourrait d'ailleurs facilement imaginer un système électrique qui permette de produire ce nombre à partir d'un complexe simplicial donné. La topologie comporte un certain nombre de transformations plus compliquées que les subdivisions barycentriques. Elles modifient l'objet topologique, c'est-à-dire donnent un objet non homéomorphe, sans changer pour autant ce qu'on appelle son « type d'homotopie ». Un invariant essentiel d'un espace topologique à homotopie près s'appelle son « groupe fondamental ». Il est trivial, c'est-à-dire réduit à un élément, dans le cadre de la sphère, mais ne l'est plus pour des complexes simpliciaux même de dimension 2, comme celui défini dans la figure 29. La topologie est l'étude des invariants des espaces topologiques, soit à homotopie, soit à

homéomorphie près. Il ne me paraît pas improbable que le cerveau utilise, peut-être sous forme élémentaire, peut-être sous forme extrêmement riche, la possibilité d'accéder à la topologie grâce à la combinatoire des complexes simpliciaux. Plus précisément, la topologie montre que la combinatoire d'un complexe simplicial peut être incroyablement élaborée. Et il serait dommage de ne pas utiliser les progrès de la topologie dans l'élaboration par exemple de machines de mémorisation, et de se restreindre aux arbres, c'est-à-dire aux complexes simpliciaux de dimension 1, dont le groupe fondamental est trivial. La définition du groupe fondamental est simple à comprendre : une fois choisi un point-base qui sert de référence, c'est-à-dire un sommet, on considère tous les chemins que l'on peut parcourir en suivant les arêtes à travers le « réseau » et qui reviennent au point de départ. On compose les chemins en les mettant bout à bout. La seule subtilité consiste à comprendre quand deux chemins définissent le même élément du groupe fondamental. Pour l'expliquer, je pourrais procéder au niveau combinatoire, mais ce serait pénible. Je peux aussi donner une image géométrique. Un complexe simplicial, bien qu'il soit un objet combinatoire, a ce qu'on appelle une réalisation géométrique. On met les sommets du complexe simplicial dans un espace de dimension suffisamment grand, on relie les sommets qui sont les extrémités d'une arête par un vrai segment, on relie les triplés qui sont les sommets d'un triangle par un vrai triangle dont les faces sont les segments précédents, et ainsi de suite. La figure ci-dessus montre un exemple de la réalisation géométrique d'un complexe simplicial de dimension 2 (*voir figure 29*). En général, il est difficile de donner une image, car le complexe simplicial est nécessairement plongé dans un espace de grande dimension. C'est pour cela qu'on ne peut le voir sous forme géométrique directe. On est donc bien obligé de remplacer la géométrie par la combinatoire. Mais, au moins en petite dimension,

il est facile d'expliquer quand deux chemins définissent le même élément du groupe fondamental. Ou de manière équivalente, quand un chemin définit l'identité comme élément du groupe fondamental. Cela a lieu lorsque l'on peut déformer le chemin pour le rendre trivial, et ce sans le déchirer (*voir figure 30*). On obtient déjà par cette construction tous les groupes intéressants à partir de complexes simpliciaux de dimension 2. C'est dire l'incroyable richesse de cette combinatoire, même pour les complexes simpliciaux de dimension 2. L'étonnant c'est que le cerveau recèle, en puissance, des myriades de possibilités de réaliser cette combinatoire et d'exploiter la richesse de la topologie. Par exemple, compter le nombre de trous d'une surface n'est autre que calculer sa caractéristique d'Euler-Poincaré.

JPC : Peut-on, sur ces bases, construire une machine? Ce serait en fait la meilleure démonstration...

AC : Compter le nombre de trous d'une surface, c'est-à-dire la caractéristique d'Euler et Poincaré, est très facile. Pour extraire l'invariant, la machine devrait compter le nombre de sommets. Elle devrait enlever le nombre d'arêtes, et ajouter celui de triangles. C'est donc extrêmement simple. Un système électrique suffirait.

JPC : Le cerveau ne procède pas ainsi. Il ne compte pas.

AC : Mais un système électrique, si. Imagine un système dans lequel chaque sommet posséderait une charge électrique égale et positive. Chaque arête posséderait une charge électrique négative. Chaque triangle ajouterait une charge positive. A ce moment, si tu regardais la charge totale, quand le système se décharge, tu obtiendrais un invariant topologique.

JPC : Il faudrait le réaliser.

AC : Bien entendu. Mais ce n'est pas impossible. Et d'ailleurs, rien n'empêche que ce ne soit pas seulement un phénomène électrique, mais aussi un phénomène chimique.

JPC : Bien sûr. J'ai retenu le phénomène électrique

parce qu'il est le plus directement mesurable. Mesurer des libérations de transmetteurs chimiques est plus difficile mais réalisable, au moins indirectement. Mais nous sommes encore loin de pouvoir le faire. La difficulté, dans le cas du système nerveux central, consiste également à repérer de petits ensembles de neurones en relation avec une dimension plus globale, plus difficile à saisir. Mais c'est possible, par exemple dans le cas de la reconnaissance des visages au niveau du cortex temporal.

AC : On pourrait penser par exemple que la reconnaissance des formes qui ne dépasse pas la topologie de dimension 2 soit effectuée par un système formé seulement de points (les neurones), d'arêtes et de triangles. C'est-à-dire un système dans lequel il ne serait pas nécessaire d'exciter, de manière corrélée, plus que des triplés de neurones. Mais évidemment tout cela relève de la pure spéculation.

JPC : Non. Voilà une prédiction simple, à soumettre aux physiologistes ! La mesure de corrélations d'activité entre neurones se fait déjà dans plusieurs laboratoires ¹. Nous avons abordé la question des invariants et des représentations. Venons-en à la deuxième question : l'organisation de la mémoire à long terme souvent représentée sous forme d'arbres... Compte tenu de la manière dont, te fondant sur cette nouvelle topologie, tu conçois les représentations, comment envisages-tu l'accès à cette mémoire et son organisation ? Mais aussi, comment peut-on raisonner par analogie ? Car le raisonnement par analogie pourrait se ramener simplement à la mise en relation de deux arbres distincts.

1. C.M. Gray, P. König, A.K. Engel and W. Singer, *Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties*. Nature n° 338, 334-337, 1989.

3. *L'organisation de la mémoire à long terme*

AC : Je sais qu'en ce qui concerne les problèmes de mémorisation, le modèle répandu est celui des arbres, dont tu viens de parler. Je ne peux éviter, sans proposer vraiment un modèle, d'expliquer une notion plus générale, plus raffinée, que la notion d'arbre : la notion de complexe simplicial *hyperbolique*, ou de courbure négative. Je n'ai pas d'idée précise pour appliquer cette notion aux procédés de mémorisation. Mais il est clair que la notion d'arbre est trop restrictive, trop rigide, et qu'elle impose, pour réparer une erreur, de faire machine arrière en suivant exactement le même chemin. La notion de complexe simplicial hyperbolique est beaucoup plus souple sans pour autant perdre les propriétés des arbres qui sont utilisées dans les modèles de mémorisation. Alors que la notion d'arbre est unidimensionnelle et organise l'information de la mémoire de manière « linéaire », les complexes simpliciaux hyperboliques, eux, l'organisent de manière plus subtile.

Qu'est-ce qu'un complexe simplicial hyperbolique ? On peut définir cette propriété de manière purement combinatoire, et dire, par exemple, qu'il est suffisant pour qu'un complexe simplicial de dimension 2 soit hyperbolique que tout sommet d'un triangle soit commun à au moins sept triangles différents. Mais on comprend beaucoup mieux la signification de cette notion en parlant de géométrie et de géodésique. Pour l'expliquer, il faut tout d'abord que je revienne aux géométries non euclidiennes. Dans le modèle de Poincaré, l'intérieur d'un disque dans le plan, les géodésiques sont les arcs de cercle perpendiculaires au bord du disque. Prenons une telle géodésique (*voir figure 31*) et un

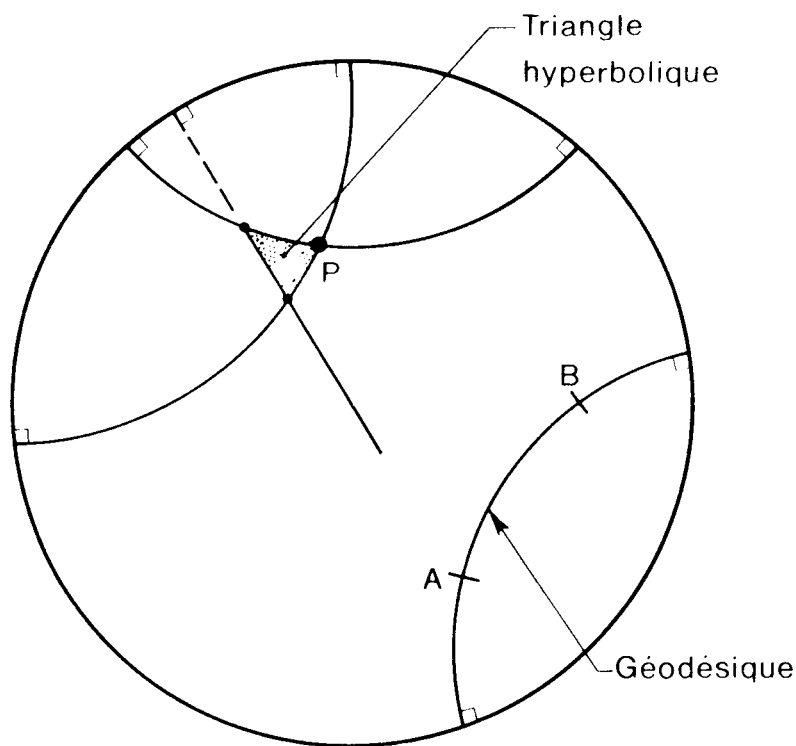


FIGURE 31

Géométrie hyperbolique. Dans cette géométrie, les points sont les points du disque ci-dessus, les droites sont les arcs de cercle perpendiculaires au bord du disque. Par le point P , situé en dehors de la droite AB , passent plusieurs droites « parallèles » à la droite AB , c'est-à-dire ne la rencontrant pas.

point P extérieur à cette géodésique. Il est facile de construire une infinité d'autres géodésiques qui passent par P et ne rencontrent pas la première. Cette géométrie où les droites sont les géodésiques ne vérifie donc pas l'axiome d'Euclide. Dans ce modèle, l'angle entre deux géodésiques est l'angle entre les cercles correspondants. On vérifie très facilement que la somme des angles d'un triangle est toujours inférieure à 180° , propriété caractéristique d'un espace de courbure négative. On peut également préciser comment mesurer la distance entre deux points dans cette géométrie de Poincaré. Le plus court chemin entre deux points A et B est la géodésique, c'est-à-dire la portion du cercle qui passe par les deux points perpendiculaires au bord du disque. Cette géométrie possède une propriété qui la rapproche d'un arbre et que ne possède pas du tout la géométrie euclidienne. Cette propriété est précisément l'hyperbolicité. Une manière simple de l'exprimer consiste à dire que si BC est un segment, et A est un point en dehors de ce segment, lorsqu'on veut se déplacer de A vers B, on ne perd que très peu, et en tout cas pas plus qu'une quantité fixée une fois pour toutes, par rapport au déplacement optimal donné par la géodésique, en la remplaçant par le chemin le plus court joignant A au segment BC suivi d'un déplacement le long de BC (*voir figure 32*). Cette propriété est évidemment vraie pour les arbres. Elle est fausse pour l'espace euclidien. Mais elle est vraie pour l'espace hyperbolique de Poincaré. De plus, j'insiste, elle continue à être vraie pour le revêtement universel d'un très grand nombre de complexes simpliciaux qui sont précisément les complexes simpliciaux hyperboliques. Pour en revenir à l'organisation de la mémoire, si l'on construisait un modèle où les objets de mémoire étaient localisés dans un espace hyperbolique, cette propriété aurait la signification suivante : pour déplacer l'attention consciente A vers un objet de mémoire X situé sur une portion finie convexe P de l'espace hyperbolique en question, il n'est pas

nécessaire de connaître à l'avance la position précise de X dans P , même si P est relativement étendue. Il suffit d'abord de se diriger vers P , puis une fois dans P de se diriger vers X . Un espace hyperbolique a exactement la propriété de cohérence nécessaire, de même que les arbres, sans être pour autant une structure unidimensionnelle, avec tous les inconvénients que cela implique.

JPC : Oui, mais entre l'avoir bien compris et l'utiliser... Il ne suffit pas de disposer d'un modèle formel théorique très général. Ce qu'il faut, pour lors, c'est faire en sorte que

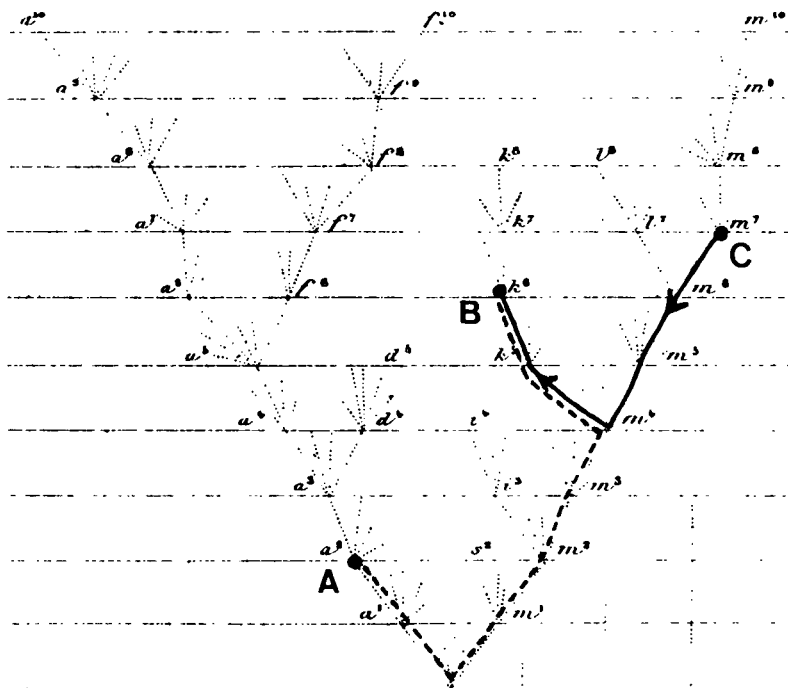


FIGURE 32

Arbre et géodésique. En gras, géodésique de C à B ; en pointillé, géodésique de A à B .

ces propositions aboutissent à des expériences réalisables en laboratoire.

AC : En fait, l'Américain W. Thurston travaille depuis quelques années sur l'utilisation de la géométrie hyperbolique pour l'amélioration des ordinateurs.

4. Le raisonnement par analogie

AC : Attends. Je n'ai pas fini de répondre à ta question. Tu me demandais aussi mon opinion sur le raisonnement par analogie.

JPC : Oui, le raisonnement par analogie arrive parfaitement à point.

AC : J'ai l'impression que le raisonnement par analogie comporte deux étapes. D'une part la reconnaissance de l'analogie, qui est probablement l'étape la plus difficile à comprendre, et qui est probablement proche de la reconnaissance des formes. D'autre part, l'étape que j'appellerai de réplication-traduction-amélioration. Le premier pas consiste à pouvoir faire la réplication d'une configuration de neurones, ou, en langage mathématique, d'un complexe simplicial qui obéit à une certaine fonction. Supposons construit un système analogue de neurones, le deuxième pas, que j'appelle « traduction », consiste alors à connecter ce système répliqué en remplaçant les mots associés au premier système de neurones par leur traduction grâce à l'analogie dans le deuxième système de neurones. La troisième phase consiste, une fois accomplie la traduction, à tester le fonctionnement du nouveau système de neurones pour améliorer sa structure. Mais le cerveau est-il capable d'effectuer une réplique ?

JPC : N'oublie pas que nos deux hémisphères sont inter-

connectés. Il n'est pas impossible que certaines représentations soient présentes simultanément sur chacun des deux hémisphères, ou qu'un transfert ait lieu d'un hémisphère à l'autre...

AC : Qu'on puisse même, dans un but déterminé, la construire, la faire passer de l'autre côté.

JPC : Des transferts considérables d'information se produisent d'un hémisphère à l'autre, mais il est encore difficile de savoir si de tels transferts interviennent dans le raisonnement par analogie. Les deux hémisphères, tu le sais, ne sont pas parfaitement symétriques. On peut imaginer qu'une représentation produite par un hémisphère soit modifiée, amplifiée ou atténuée par l'autre. Comprendre les relations entre hémisphère droit et hémisphère gauche est un problème très important. Chez des mammifères inférieurs, qui ne disposent pas du langage, il n'existe que très peu, voire pas de latéralisation. Celle-ci semble apparaître avec le développement du langage. Ce qui permet, en plus, d'utiliser les deux hémisphères de manière indépendante, et donc de créer un accroissement « explosif » de la surface du cortex utilisable désormais de manière non redondante. Des changements géniques mineurs, qui créent une légère asymétrie dans le développement, ont pu produire brutalement, du fait de l'épigenèse, une suppression brutale de redondance et une exploitation de possibilités d'un hémisphère par l'autre et réciproquement. C'est peut-être là l'origine du « phénomène humain », qui fait que sur la base d'un petit nombre de gènes modifiés et d'un accroissement du cerveau qui n'est pas, somme toute, gigantesque, on accède à un niveau de performance d'un nouvel ordre.

5. Enchaînements de représentations et cadres de pensée

JPC : Peut-être pouvons-nous revenir à l'idée de départ : le darwinisme en mathématiques, l'enchaînement, pour construire un raisonnement, et la « ruminantion », avant l'illumination, d'objets mathématiques confrontés les uns avec les autres dans le cadre d'un problème défini. Distinguons, pour simplifier, deux questions. 1° L'enchaînement dans le temps de représentations mentales conduit à une « proposition » sous une forme falsifiable, elle-même exprimant une « vérité ». 2° La définition de ce qu'on appelle un « cadre de pensée » ou « intention » qui se trouve à la base de la réflexion et même de la création mathématique. Comment définir une intention en mathématiques ?

AC : La théorie des probabilités contient un concept important, celui de « conditionnement », qui pourrait peut-être s'appliquer ici. Il me semble nécessaire, pour définir une intention, comme, par exemple, celle de gagner une partie d'échecs, de l'identifier à une fonction d'évaluation qui sert à estimer à quelle distance on se trouve du but à atteindre. Il reste pourtant à comprendre comment le cerveau construit cette fonction d'évaluation. Nous y reviendrons. Pour l'instant, supposons que nous disposons de cette fonction d'évaluation et utilisons-la, comme dans la théorie des probabilités, pour conditionner les systèmes. J'ai alors envie d'employer l'image suivante, qui correspondrait au darwinisme dont tu parles : en nous fondant sur des mécanismes d'évolution interne comme l'analogie, supposons que le cerveau ait déjà élaboré un millier d'assemblées de neurones, de complexes simpliciaux de neurones, et qu'il les

mette en activité en les conditionnant par la fonction d'évaluation. Chaque système produit un résultat et il faut que le cerveau puisse sélectionner parmi ces résultats celui qui optimise la fonction d'évaluation. Je pense que les physiciens ont trouvé, grâce au principe de la phase stationnaire, une très bonne idée pour sinon résoudre ce problème, du moins suggérer un mécanisme intéressant. Supposons que chacun des systèmes de neurones produise un courant électrique dont la phase est proportionnelle à la valeur de la fonction d'évaluation sur ce système. Dans le cas de systèmes qui ne réalisent pas le maximum de cette fonction, l'existence de valeurs voisines, inférieures et supérieures, de la fonction pour d'autres systèmes, entraîne alors l'annulation de la somme des courants produits. Dans le cas des systèmes dans lesquels la fonction d'évaluation est maximale, cette annulation n'a pas lieu. Ils sont donc les seuls à contribuer de manière sensible au courant résultant de tous les courants partiels. Ce genre de système n'est certes pas économique, et on peut en imaginer de beaucoup plus simples lorsque la fonction d'évaluation est spécifiée une fois pour toutes, comme dans les ordinateurs qui jouent aux échecs. Il n'empêche qu'il a une très grande flexibilité, que les physiciens exploitent constamment grâce à l'intégrale de Feymann.

JPC : On dispose donc presque d'un mécanisme de sélection.

AC : Oui, mais malheureusement, à condition d'avoir déjà construit la fonction d'évaluation. Comment ? J'avoue n'en avoir aucune idée, même très vague.

JPC : Il le faudrait pourtant.

AC : Je pense seulement que cette fonction doit être liée au système limbique, ou à d'autres parties du cerveau. Elle ne peut être purement interne.

JPC : C'est déjà une idée que d'imaginer qu'une boucle puisse exister.

AC : Je pense qu'une corrélation doit exister entre la

fonction d'évaluation et la frustration ou le plaisir qu'on peut ressentir quand on est proche de la résolution d'un problème. Mais je ne sais pas exactement comment la définir. Comment le but spécifique, qui est celui de la pensée active, de la pensée présente, peut-il se manifester dans le mécanisme d'interférences destructives ou constructives?

JPC : On peut imaginer que lorsque le but est atteint, à ce moment-là...

AC : Oui, mais cela ne me satisfait pas, parce qu'en théorie des probabilités, avant ce phénomène d'interférences destructives ou constructives, la probabilité est tout de même conditionnée, et ce à l'avance, par le but que l'on doit atteindre.

JPC : Nous y voilà. Le calcul s'effectue dans un cadre! C'est très important. Tu ne veux pas distinguer, comme moi, le raisonnement logique et...

AC : Je ne parle pas de la création. J'en reste au deuxième niveau.

JPC : Même à ce deuxième niveau, le but à atteindre joue un rôle. Il est fixé. On pourrait même le considérer comme une « obsession » interne. Cela veut dire qu'un état persistant d'activité des neurones...

AC : ...devrait engendrer une espèce de frustration ou de gêne.

JPC : Attends. On peut aussi imaginer qu'un circuit, bouclé sur lui-même, fait intervenir le système limbique parce qu'il y a désir. Le cerveau produit une hypothèse de plaisir qui joue le rôle de guide, et ouvre l'accès à une solution qui est ou non source de plaisir...

AC : Ou inversement, parce qu'il y a frustration. On le ressent très souvent en mathématiques. Quand quelque chose cloche, ce n'est pas un désir, c'est une frustration qui intervient.

JPC : Par inquiétude de ne pas avoir atteint le but. Le

système limbique maintient « en activité » une représentation qui crée un contexte dans lequel vont s'ajuster d'autres représentations mentales qui, en fin de compte, entreront en résonance avec la fin qu'on se propose. Elles procurent alors un sentiment de plaisir, d'« achèvement » de la représentation initiale. C'est une métaphore, mais sur des bases semblables, Stanislas Dehaene et moi-même ¹ avons construit récemment un modèle d'apprentissage de « règles » qui semble fonctionner.

AC : Je suis d'accord, mais ton image ne rend pas bien compte de la possibilité de mesurer la proximité du but. Tant qu'il n'est pas atteint, il faut, pour que l'image soit valide, que sa *proximité* puisse être reconnue. Même s'il n'est pas atteint. C'est essentiel pour que le conditionnement se produise. J'admets bien la possibilité de reconnaître que le but est atteint. Par contre, introduire une distance par rapport au but, c'est-à-dire pouvoir tout conditionner, me paraît plus difficile...

JPC : Peut-être un progrès dans l'actualisation de l'intention l'enrichit-il progressivement par accumulation.

AC : On touche ici un point très important pour la pratique des mathématiques. Il arrive souvent, dans un problème, qu'évaluer à quelle distance se trouve la solution facilite la tâche. C'est précisément cette intuition grossière du chemin qui reste à parcourir qui, même si les questions étudiées peuvent paraître très bizarres, aide à résoudre le problème.

1. S. Dehaene, J.-P. Changeux, *A model of prefrontal cortex*, Cognitive neuroscience, 1989 (sous presse).

6. La sélection naturelle des objets mathématiques

JPC : As-tu d'autres remarques à faire sur le darwinisme en mathématique ?

AC : Je considère que le darwinisme du fonctionnement cérébral repose sur des mécanismes d'interférences constructives et de résonances de groupes, plutôt que sur un phénomène de sélection naturelle ou d'élimination.

JPC : Je pense que c'est une forme de sélection naturelle. Mais « sélection naturelle » est à prendre ici en un sens précis et conforme à ce qu'on connaît de la structure et du développement du cerveau. Cette notion, même en dynamique des populations, est souvent difficile à préciser. Elle est définie dans des termes qui sont ceux de populations qui se reproduisent selon une répartition géographique définie. Le darwinisme traditionnel, appliqué à l'évolution des espèces, contient les notions de dynamique temporelle, de population et de répartition géographique. Dans le cas du système nerveux, la composante de multiplication n'est pas utilisable. Les neurones ne se multiplient pas. Seule compte l'occupation différentielle et « compétitive » de certains territoires. Ta formulation va tout à fait dans ce sens. Les mécanismes d'interférences constructives et de résonances de groupes peuvent alors être considérés comme des mécanismes de sélection propres au cerveau.

Venons-en au troisième niveau. Comment vois-tu les intentions ?

AC : La caractéristique fondamentale de ce niveau, dans l'illumination, c'est, au-delà du plaisir ressenti, l'impression tout à coup qu'un brouillard se lève brutalement. La fraction

consciente de la pensée accède alors directement à un monde dépourvu pour elle de toute étrangeté. Nulle vérification laborieuse n'est plus nécessaire. C'est sans doute cette sensation, caractéristique du troisième niveau, qui excite le système limbique.

JPC : Tu me fais penser à l'extase mystique de sainte Thérèse d'Avila.

AC : L'extase mystique doit certainement exciter les mêmes régions du cerveau. Mais pour d'autres raisons. De même que l'harmonie esthétique.

JPC : Nous touchons ici une question qui me tient à cœur, celle des relations entre la science et l'art. Quelle différence y a-t-il entre l'objet mathématique et l'œuvre d'art ?

AC : Il n'est pas impossible que les artistes, poètes ou musiciens, parviennent, avec leurs propres ressources, à exprimer des données extrêmement élaborées qui témoignent de l'harmonie que l'on ressent, peut-être une fois dans une vie, à travers l'illumination. Il arrive en effet qu'une œuvre d'art, comme un morceau de musique, excite le système limbique de manière tout à fait semblable. Mais pour en revenir à l'illumination et aux mathématiques, comme le seul moyen que nous avons de transmettre un résultat est une chaîne logique de raisonnements, il faut bien revenir rapidement du troisième niveau au premier. Il est nécessaire de procéder à une vérification pédestre de la démonstration que l'illumination a permis d'entrevoir. L'état d'excitation est donc extrêmement bref. Une fois décomposée, la démonstration peut se vérifier pas à pas. Mais en un sens, la phase purement « mystique », elle, s'est évaporée.

JPC : Comment le darwinisme intervient-il ?

AC : Je pense que la première des trois étapes dont parle Hadamard, c'est-à-dire la préparation, permet précisément de définir la fonction d'évaluation qui doit, en principe, conditionner le darwinisme.

JPC : Le darwinisme en évolution part de l'amibe pour arriver à l'homme. C'est là son intérêt. Son application la plus intéressante aux mathématiques est la « création » d'un nouvel objet mathématique par combinatoire des éléments qui font déjà partie des mathématiques établies. Il y a genèse d'un « monstre », d'une « chimère ». Cet objet, bien que nouveau, est éventuellement sélectionné en vertu de sa résonance avec un corpus qui existe déjà.

AC : On peut très bien envisager cela.

JPC : A l'appui de cette idée, je rappellerai que, selon toi, dans l'acte créateur, aborder un problème consiste d'abord à « l'élargir ». Que veut dire l'élargir ? C'est faire entrer dans le compartiment de travail de la mémoire à court-terme des objets mathématiques dépourvus de lien direct avec le but fixé. L'intrusion d'étrangers, d'« outsiders », donne naissance à un nouvel objet mathématique. Elle permet de casser le cadre dans lequel le mathématicien se trouvait auparavant, et lui ouvre l'accès à un nouveau niveau de connaissance. Celui-ci est alors la résultante de la combinatoire à l'œuvre, parfois très longuement, durant la période d'incubation. Le darwinisme en mathématique paraît rendre compte de la « création » de manière particulièrement adéquate. La balle est dans ton camp. Quelles conditions de sélection produisent l'illumination ? Est-ce l'intégration à tout ce qui existait auparavant ?

AC : Je ne sais pas si l'on peut conserver l'image d'une interférence constructive, telle que je l'ai adoptée précédemment. L'illumination, lorsqu'elle se produit, ne porte pas seulement sur l'objet en question, pris dans sa nouveauté, mais aussi sur sa cohérence avec ce que le cerveau a déjà compris et connaît bien.

JPC : Et les différences ? On ne retombe pas sur une structure déjà connue. Il n'y a donc pas simple conformité. L'objet est nouveau, et néanmoins s'intègre à tout ce que l'on connaît déjà.

AC : Je ne sais pas comment dire. On n'a plus besoin d'un mécanisme d'évaluation en fonction d'un but déterminé, mais d'une mesure immédiate de cette compatibilité, avant même que la pensée réfléchie n'entre en jeu. Un mécanisme, malaisé à comprendre, permet, sans recours à la pensée raisonnée, de ressentir la résonance entre le nouvel objet de pensée et ceux qu'on est habitué depuis longtemps à manipuler. Tout cela est, je l'avoue, très difficile à comprendre...

JPC : Oui, mais une machine qui soit créatrice en mathématiques nécessiterait ces mécanismes.

AC : Exactement. Sinon, ce serait un ordinateur traditionnel. Il est remarquable que le cerveau puisse percevoir cette cohérence entre des objets différents, ainsi que l'harmonie d'un objet qu'il ne connaissait pas auparavant. Mais il n'y a pas du tout identité. C'est là, selon moi, la cohérence du monde mathématique.

JPC : C'est la congruence avec d'autres objets mathématiques qui se trouvent dans la mémoire à long terme.

AC : Il me semble seulement que c'est la preuve de la cohérence du monde mathématique, indépendamment de l'individu.

JPC : C'est là où je voulais t'amener. Cette cohérence intervient dans le processus de sélection d'abord sur la base de la non contradiction et une nouvelle cohérence en résulte.

AC : Je n'en suis pas certain. Je crois seulement qu'elle *se manifeste à travers* ce processus de sélection...

JPC : Ne revenons pas sur ce débat ! Je pense que l'intégration d'un objet nouveau dans un ensemble ouvre un nouvel espace de connaissance... L'illumination, en quelque sorte, met en harmonie plusieurs niveaux d'organisation du cerveau, comme dans le cas de la contemplation de l'œuvre d'art. Mais comment définir cette forme de joie esthétique que certains tableaux nous procurent ¹ ? Elle semble s'ex-

1. J.-P. Changeux, *Raison et plaisir*, loc. cit., p. 158.

pliquer par de multiples résonances entre divers niveaux liés à la fois à la rationalité, à l'entendement et au système limbique. L'entrée en résonance se produit lorsque le spectateur se trouve en face d'une structure « singulière ». On peut donc concevoir cette illumination comme une espèce d'objet mental inter-niveaux, nouveau par rapport à tout ce qui a pu exister, et qui met en relation des objets mentaux disjoints.

AC : Je suis tout à fait d'accord avec ton interprétation. J'aimerais pourtant que tu la précises.

JPC : Cette métaphore s'applique tant à l'œuvre d'art que, dans une certaine mesure, à l'illumination mathématique. L'illumination est d'autant plus forte que l'objet qui surgit est nouveau et envahit un domaine occupé déjà par des structures latentes. Tu prétends, toi, que ces structures sont là *pour* qu'il vienne...

AC : Oui, mais elles ne sont pas internes au cerveau. Elles sont propres au monde mathématique.

JPC : Que les mathématiques existent ou non dans le monde extérieur, elles sont dans le cerveau au moment où l'illumination a lieu.

AC : Tout à fait. Je voulais seulement dire, pour en revenir à notre discussion du début, qu'après avoir fait l'expérience de l'illumination, il est difficile de ne pas croire en l'existence d'une harmonie indépendante du cerveau et qui ne doit rien à la création individuelle.

JPC : C'est subjectif. Je crois qu'on ne peut pas dire que la *Pietà* de Michel-Ange existe avant que Michel-Ange l'ait faite. Reprenons une fois de plus la métaphore artistique. On ressent une « illumination » quand on voit pour la première fois le Jugement dernier. Mais il est absurde de dire que ce tableau existait avant que Michel-Ange ne l'ait peint. De même en mathématiques...

AC : Il y a sûrement du vrai dans ce que tu dis. Je pense tout de même qu'il existe une différence fondamentale

entre l'harmonie que l'on ressent devant la *Pietà* de Michel-Ange et celle que l'on ressent, par une belle nuit d'été, en vérifiant à l'aide d'une lunette astronomique et d'une calculette que les quatre satellites de Jupiter tournent autour de cette planète en obéissant aux lois de Kepler. Il me paraît bien difficile d'admettre que ce genre d'harmonie cosmique soit une production du cerveau humain. J'irai même jusqu'à dire, au contraire, que cette harmonie préétablie, bien antérieure à l'homme, a probablement contribué, à travers « la mystérieuse profondeur des nuits étoilées », à susciter la curiosité métaphysique. Mais revenons à l'illumination.

JPC : Encore une fois, ne confondons pas l'existence de régularités dans le monde matériel et leur expression, en termes approximatifs, par des équations mathématiques, produits du cerveau de l'homme. Pour progresser sur le plan théorique, il serait intéressant de définir – c'est peut-être un des grands avantages du modèle darwinien – les propriétés du *générateur de diversité* et son mode de fonctionnement dans la fixation des intentions, le recrutement d'objets de mémoire dans le compartiment à long terme, et surtout ses *critères de sélection*. Comment conçois-tu cette fonction de sélection ? Si ce n'est pas une fonction d'évaluation, de quoi s'agit-il ? Il y a quand même évaluation d'une cohérence qu'il faut vérifier, valider... N'est-ce pas la plausibilité d'une hypothèse ?

AC : Il est frappant que cette évaluation de la cohérence mathématique se produise de manière instantanée. En une fraction de seconde, apparaît non seulement la plausibilité, mais aussi la certitude de l'adéquation de ce qu'on a trouvé avec ce qu'on cherche. Ce n'est pas un réflexe, mais cela se produit à la même vitesse.

JPC : On retrouve la reconnaissance des visages. Pas celle d'un visage connu. Mais d'un inconnu qui convient.

AC : C'est précisément, je crois, ce qui distingue le deuxième niveau du troisième. Le deuxième est capable de

reconnaître ce qui résout un problème fixé à l'avance avec les moyens élaborés par une stratégie. Le troisième, lui, est capable de comprendre l'harmonie et la puissance d'un objet nouveau qui ne correspond pas forcément à un problème spécifique.

VI

Les machines à penser

1. Des machines intelligentes?

JEAN-PIERRE CHANGEUX : Le titre même de cet entretien pose le problème essentiel de la relation du cerveau avec la machine et, plus généralement, celle des sciences exactes avec le cerveau et son fonctionnement. Dans le domaine des machines à penser, on peut distinguer au moins trois approches.

La première, celle de l'*intelligence artificielle*, se propose de simuler les fonctions supérieures du cerveau, de l'intelligence humaine, à l'aide d'un ordinateur. Il s'agit en quelque sorte de remplacer le cerveau de l'homme par une machine. Les réussites de l'intelligence artificielle sont nombreuses : des robots qui peignent des voitures, des ordinateurs qui commandent les voyages de vaisseaux spatiaux jusqu'à Mars et au-delà, des systèmes experts qui renseignent sur les progrès les plus récents de la médecine, etc. Toutefois, l'intelligence artificielle n'ambitionne pas de comprendre

comment le cerveau humain fonctionne, mais seulement de « simuler » certaines de ses fonctions. Cette démarche est donc d'emblée très limitée.

La deuxième approche s'efforce de *modéliser* le cerveau humain et ses fonctions. Il s'agit d'un travail de recherche plus profond, qui suppose la contribution multidisciplinaire de compétences issues des mathématiques, de la physique, de la neurobiologie et de la psychologie. Cette modélisation exploite des données d'ordre anatomique et physiologique, les résultats de la biologie moléculaire et, bien entendu, l'observation du comportement qui relève de la psychologie et de l'éthologie. Elle est encore peu avancée. On dispose pourtant d'assez bons modèles en ce qui concerne quelques mécanismes élémentaires comme la propagation de l'influx nerveux (modèle de Hodgkin et Huxley), ou les transitions allostériques de récepteurs postsynaptiques, mais aussi des systèmes composés de peu de cellules nerveuses, comme ceux qui interviennent dans la nage de la lamproie ¹, la capture d'une information visuelle par une rétine artificielle, ou enfin, l'apprentissage du chant chez les oiseaux ². Je pense que cette approche est de loin la plus importante. Et si nous discutons ensemble de ce thème, c'est peut-être parce que nous pouvons l'un et l'autre apporter notre contribution dans ce domaine.

Venons-en à la troisième approche, celle des *machines neuro-mimétiques*. Le projet est le suivant : une fois des modèles théoriques des fonctions cérébrales établis à partir de l'objet naturel que constitue le cerveau et ses neurones, il s'agit de construire des machines capables, sur la base d'architectures neurales réelles, de conduites authentiquement intelligentes.

1. S. Grillner, P. Wallen, N. Dale, L. Brodin, J. Buchanan & R. Hill, *Transmitters, membrane properties and networks circuitry in the control of locomotion in lamprey*, Trends Neurosc. n° 10, 1987, p. 34-41.

2. Dehaene, Changeux et Nadal, loc. cit., p. 154.

Trois approches, donc, mais encore très peu de résultats. Les architectures employées demeurent très simplistes : quelques couches de cellules nerveuses, des mécanismes élémentaires rudimentaires, etc.

ALAIN CONNES : La deuxième ne va-t-elle pas sans la troisième ?

JPC : Oui. La troisième approche est en quelque sorte une vérification de la deuxième. Pour montrer qu'un modèle théorique est adéquat, il faut l'« expérimenter » en construisant une machine dont les performances ressemblent à celles du cerveau humain. On peut donc considérer qu'effectivement cette troisième approche complète la deuxième.

Mais je souhaiterais qu'ensemble nous abordions maintenant trois questions. La première concerne le théorème de Gödel, la deuxième, la machine de Turing, et la dernière, les différences ou les ressemblances entre le cerveau humain et les machines qu'aujourd'hui il parvient à construire lui-même.

2. Le théorème de Gödel

JPC : Dans les ouvrages qui traitent de biologie, le théorème de Gödel est fréquemment invoqué pour modérer les ambitions des neurobiologistes, ou même pour mettre en question leur approche. Il sert alors à justifier l'idée selon laquelle « l'esprit humain » résistera toujours à la science. François Jacob écrit par exemple : « On peut être assuré que les réactions caractérisant l'activité du cerveau apparaîtront aux biochimistes toutes aussi banales que celles de la digestion, mais décrire en termes de physique et de chimie un mouvement de la conscience, un sentiment, une décision, un souvenir, c'est là une autre affaire. Rien ne dit qu'on y

parvienne jamais, pas seulement à cause de la complexité, mais aussi parce qu'on sait depuis Gödel qu'un système logique ne peut suffire à sa propre description¹. » A l'opposé, on connaît le célèbre aphorisme de Cabanis : « Le cerveau sécrète la pensée comme le foie la bile. » Pour ma part, je partage le point de vue de François Jacob sur la biochimie du cerveau et sur le caractère relativement banal des molécules qui contribuent à la structure et aux fonctions élémentaires de notre cerveau. Les données qui ont été obtenues depuis 1970 le montrent. Mais je n'adhère pas à sa manière d'appliquer le théorème de Gödel aux neurosciences. Un problème méthodologique intéressant pourrait, certes, se poser, si le neurobiologiste examinait son propre cerveau en train de s'examiner. Néanmoins, je ne vois pas, dans l'état actuel de la science, d'obstacle fondamental à l'étude du fonctionnement cérébral d'un collègue, ou du tien par exemple, à l'aide de méthodes d'imageries non invasives, ou mieux de celui d'une espèce animale proche de l'homme, comme le singe, par les méthodes de la neurophysiologie expérimentale. Pour la bonne raison que la méthode dite de réduction, ou mieux encore de reconstruction, que nous employons tous dans les sciences expérimentales, consiste à rechercher l'explication à un niveau inférieur à celui que l'on veut expliquer. On se fonde alors sur l'organisation, les règles d'interaction et les propriétés des éléments qui composent le niveau inférieur pour expliquer les propriétés du niveau supérieur. Le neurobiologiste recherche ainsi les bases neurales des fonctions supérieures du cerveau humain. Et, à ce stade, de mon point de vue, aucun obstacle théorique ne se dresse. Les obstacles majeurs me paraissent bien plus la complexité dans l'organisation du cerveau, sa variabilité d'un individu à l'autre, et l'éventuelle interférence des moyens d'observation avec le fonc-

1. *La logique du vivant*, p. 337, Paris, 1970.

tionnement cérébral lui-même. On retrouve d'ailleurs ce problème en physique, où les méthodes d'observation peuvent interférer elles aussi avec les objets observés.

Revenons au théorème de Gödel. On peut considérer que sa traduction mathématique est contenue dans le célèbre paradoxe philosophique : « Tous les Crétois sont des menteurs, dit Epiménide, penseur crétois. » Il est impossible de décider si cette assertion est vraie ou fausse. On se trouve donc dans une situation d'indécidabilité. Quelle est ta définition du théorème de Gödel ? Comment l'appliquerais-tu aux neurosciences et, plus particulièrement, à la modélisation du fonctionnement cérébral, du cerveau produisant des mathématiques ?

AC : A ma connaissance, il existe deux résultats fondamentaux de Gödel en ce qui concerne l'impossibilité, comme disait F. Jacob, pour un système logique de suffire à sa propre description. Le premier indique qu'il est impossible, du fait d'un mécanisme d'autoréférence, de démontrer que la théorie des ensembles est non contradictoire. C'est d'ailleurs vrai pour toute théorie, même plus rudimentaire, pourvu qu'elle contienne certains axiomes très simples. Vient ensuite le théorème d'incomplétude. Pour expliquer ce deuxième résultat, il me faut d'abord préciser ce qu'est, dans un système d'axiomes comme celui de la théorie des ensembles, une proposition indécidable. Je voudrais, pour l'expliquer, raconter une petite histoire. Plusieurs années durant, je suis allé, chaque jeudi, chez un ami mathématicien qui pensait avoir démontré un théorème. Il travaillait sur le problème, qui porte le nom d'un mathématicien polonais d'avant-guerre, de savoir si l'ensemble ordonné des nombres réels est caractérisé par une certaine propriété. Pendant presque trente ans ce problème a occupé mon ami. Et chaque jeudi, quand j'allais le voir, il proposait une solution. Il pensait avoir obtenu une démonstration, et, chaque fois, nous procédions de la même manière. Il me

proposait sa solution, souvent sous forme écrite. Je cherchais l'erreur. Parfois je trouvais immédiatement, parfois il nous fallait en reparler la semaine suivante. Et chaque fois il recommençait, modifiait sa démonstration, encore et encore. En fait, dès le début, je savais que toute démonstration était impossible. Mais je savais aussi qu'il m'était impossible de lui montrer son erreur en lui donnant un contre-exemple. Pourquoi? Parce qu'il avait été démontré, dans les années soixante, que ce problème est indécidable. Il en est parfois ainsi en mathématiques. Dans ce cas précis, on sait que si, aux axiomes de la théorie des ensembles, on en ajoute un autre, par exemple celui du continu, on pourra démontrer que la réponse au problème est oui. Mais si on ajoute un autre axiome convenable, on pourra alors démontrer que la réponse est non. En d'autres termes, la situation est telle qu'il est impossible à un mathématicien, sans ajouter d'autres axiomes à ceux de la théorie des ensembles, de démontrer le résultat. Mais il m'était également impossible de lui donner un contre-exemple sans utiliser un axiome additionnel auquel il lui était facile de s'opposer. Il faut bien comprendre ce qu'est l'indécidabilité. Elle a toujours un sens...

JPC : ...au sein d'un système donné d'axiomes.

AC : Exactement. Un énoncé est indécidable, si l'on peut ajouter soit sa véracité, soit sa fausseté, sans contredire les axiomes avec lesquels on travaille quotidiennement, au-delà d'une contradiction possible de la théorie des ensembles.

JPC : Les axiomes internes au système ne suffisent donc pas à la décision.

AC : Oui. Nous pouvons maintenant énoncer le théorème d'incomplétude de Gödel. Il dit que quels que soient les axiomes en nombre fini ou donnés de manière récursive que l'on prend, il y a toujours des questions auxquelles on ne peut répondre, qui restent indécidables, et pour lesquelles on aura encore un manque d'informations. Autrement dit, le théorème de Gödel précise qu'il est impossible de prendre

un nombre fini d'axiomes de telle sorte que toute question soit décidable. Cela ne veut pas dire qu'on ne puisse pas analyser une question à partir de ce que l'on sait, mais cela signifie que le nombre de questions intéressantes et nouvelles dont il faudra ajouter la réponse est infini. Voilà comment il faut comprendre le théorème de Gödel. Ce serait, je pense, une erreur d'en déduire que le pouvoir de la machine humaine est limité. Ce théorème dit seulement qu'avec un nombre fini d'axiomes, on ne peut pas avoir réponse à tout. Mais si une question n'est pas décidable, à condition de l'avoir démontré, on peut lui attribuer une réponse et continuer à raisonner.

Cela signifie que chaque nouvelle question indécidable donne lieu à une bifurcation selon que l'on choisit la réponse positive ou négative. Le monde dans lequel on se promène comporte plusieurs bifurcations possibles. C'est absolument tout ce que cela signifie. Une fois qu'on a attribué une réponse à une question, on peut continuer et se poser de nouvelles questions. D'anciennes questions deviennent alors décidables qui ne l'étaient pas auparavant... Chaque question indécidable crée une bifurcation et impose un choix. Par exemple, au théorème de Paul Cohen sur l'hypothèse du continu succède une bifurcation : on choisit ou bien qu'il n'y ait pas de cardinaux entre le dénombrable et le continu, ou bien qu'il y en ait trente-six. La première réponse s'impose par sa simplicité. Mais il est très important que les choix de réponse portent sur les questions les plus primitives possible, et il existe en fait des questions plus primitives que celle du continu.

JPC : Tu ne vois pas d'obstacle théorique fondamental...

AC : Je ne parle pour le moment que du problème d'indécidabilité. En présence de questions indécidables comme celle du continu, il faut arriver à formuler une hypothèse qui les rende décidables, puis à tester les conséquences de cette hypothèse, et sa capacité à éclairer d'autres

questions. Par exemple, si l'on utilise l'hypothèse du continu, on peut démontrer – c'est là un résultat de G. Mokobodski – qu'il est possible d'attribuer une limite $\text{Lim}_w (a_n)$ à toute suite bornée de nombres réels de telle sorte que $\text{Limite inférieure } (a_n) \leq \text{Lim}_w (a_n) \leq \text{Limite supérieure } (a_n)$ et que Lim_w dépende mesurablement de (a_n) et commute avec l'intégrale. C'est là un résultat bien utile dans les mathématiques que j'utilise. Quand on ajoute une hypothèse comme celle du continu, il faut évidemment s'assurer de son indépendabilité, c'est-à-dire de deux choses : d'une part, elle ne doit pas résulter des axiomes précédents – c'est là un théorème de P. Cohen pour l'hypothèse du continu – et, d'autre part, sa négation, elle, ne doit pas résulter des axiomes précédents, ce qui, pour l'hypothèse du continu, est un résultat de K. Gödel. En fait, on démontre toujours ces résultats en supposant que la théorie des ensembles est non contradictoire. Mais je considère qu'il n'est pas pertinent d'utiliser le théorème d'incomplétude pour limiter notre mécanisme de compréhension. Il faut simplement comprendre qu'il y aura des choix à faire et qu'on ne pourra donner un procédé récursif pour les faire une fois pour toutes. Voilà ce que signifie ce théorème.

JPC : La réponse est une alternative. Ce théorème porte plutôt sur le processus d'acquisition de la connaissance que sur une impossibilité logique ou épistémologique. Les neurobiologistes peuvent donc avoir confiance. Tôt ou tard, nous comprendrons le fonctionnement du cerveau !

AC : Ce théorème définit une sorte d'horizon de compréhension déterminé par le nombre fini des choix déjà effectués. Plus ce nombre est grand, plus l'horizon est éloigné. Il ne faut pas avoir une vision statique selon laquelle il devrait exister, une fois pour toutes, un nombre fini d'axiomes donnant réponse à tout. Notre compréhension est au contraire dynamique. Chaque fois qu'elle s'accroît, nous pouvons donner des réponses à de plus en plus de questions,

nous pouvons faire des choix à chaque nouvelle bifurcation, de sorte que notre horizon s'éloigne. Il est évidemment illusoire de penser qu'un jour, nous aurons tout compris. C'est le problème de la science en général. Mais il ne faut pas se laisser limiter et décourager par l'énoncé de ce théorème.

En fait, dans sa formulation la plus profonde, le théorème d'incomplétude de Gödel montre que l'on ne peut réduire les mathématiques à un langage formel. Au début du siècle, les mathématiciens ont cherché à préciser ce qu'est une démonstration en mathématiques. Hilbert a construit un langage artificiel fondé sur un alphabet fini, un nombre fini de règles de grammaire permettant de spécifier sans ambiguïté quelles sont les propositions cohérentes, et un nombre fini de règles d'inférence logique et de propositions supposées vraies ou axiomes. A partir d'un tel système ou langage formel, un algorithme universel permet de décider de la validité d'une démonstration formulée dans ce langage. On peut ainsi, au moins théoriquement, établir la liste de tous les théorèmes démontrables dans ce langage formel. Hilbert espérait pouvoir réduire les théorèmes mathématiques à ceux qui sont démontrables dans un langage formel convenable. Le théorème de Gödel montre que c'est impossible. Quelle que soit la complexité d'un système formel, il y aura toujours un énoncé, portant sur les nombres entiers positifs, qui sera à la fois vrai et indémontrable dans le système formel. On a beaucoup insisté sur l'aspect négatif de ce théorème qui rend impossible une définition claire de ce qu'est une démonstration. Mais ne peut-on l'envisager sous l'angle suivant : les propositions vraies portant sur les nombres entiers positifs ne peuvent se réduire, via inférence logique, à un nombre fini d'axiomes. La quantité d'information contenue dans l'ensemble de toutes ces propositions est donc infinie. N'est-ce pas là une caractéristique d'une réalité indépendante de toute création humaine ?

Mais venons-en au problème de l'introspection. Dès le début de la théorie des ensembles, des paradoxes, comme celui de Russel, obligent à hiérarchiser les propositions logiques selon des types successifs. Ce paradoxe apparaît dès que l'on fait des fautes de syntaxe. Par exemple, si l'ensemble de tous les ensembles formait un ensemble, on pourrait en considérer une partie, qui serait l'ensemble des ensembles qui se contiennent comme éléments, qui aurait comme complémentaire l'ensemble des ensembles qui ne se contiennent pas comme éléments. Le paradoxe apparaît lorsqu'on se demande si cet ensemble se contient ou ne se contient pas comme élément. Pour y répondre, il suffit de hiérarchiser, en logique, les éléments selon un *type* différent de celui des ensembles. On commence par les éléments, de type 0. Ensuite viennent les ensembles, de type 1. Cette distinction entre des types d'échelons différents permet de ne pas les mélanger. Il devient donc impossible de parler de l'ensemble de tous les ensembles, notion qui représente une faute de syntaxe. Quand la logique est hiérarchisée, ce paradoxe disparaît.

JPC : Il s'agit en quelque sorte de créer un ordre.

AC : La succession des types permet une hiérarchie dans les mécanismes de pensée : on considère les éléments comme plus simples, moins élevés, que les ensembles.

JPC : On ne peut raisonner dans les deux sens.

AC : On ne peut mettre sur le même plan un élément et un ensemble. En particulier, on ne peut poser la question de l'ensemble des ensembles qui se contiennent comme éléments. Un procédé analogue doit pouvoir s'appliquer au problème de l'introspection du cerveau se comprenant lui-même et éliminer par là même ce prétendu paradoxe.

JPC : Donc il n'est pas indécidable.

AC : Il n'est pas question ici d'indécidabilité. Le paradoxe résulte d'une faute de syntaxe. On a compris qu'il fallait formuler la logique de la théorie des ensembles de

telle sorte que ce paradoxe soit éliminé. Et il l'est, à partir du moment où les questions sont formulées compte tenu de cette hiérarchie.

JPC : Tu rends le paradoxe décidable en lui ajoutant des hypothèses.

AC : C'est un paradoxe qui force à se donner une notion plus raffinée des objets logiques et à introduire une hiérarchie entre eux.

JPC : Passons à la deuxième question.

AC : Oui. En quel sens le théorème de Gödel impose-t-il une limite à la compréhension du fonctionnement cérébral? Les mathématiciens¹ ont compris, en analysant la notion de suite aléatoire, le lien étroit qui existe entre le théorème de Gödel et la théorie de l'information inventée au début des années cinquante. Au point que l'on peut considérer ce théorème comme une conséquence des contraintes imposées par la théorie de l'information, à cause de la finitude de la complexité de tout système formel. Ainsi, des deux limitations avancées par F. Jacob, la complexité et le théorème de Gödel, la deuxième est la conséquence de la première. On peut donc formuler une parade à ces limitations. On introduit d'abord, pour éviter le paradoxe d'introspection, une hiérarchie entre cerveaux analysés (type 0) et cerveaux analysants (type 1)... De plus, en utilisant le caractère évolutif de la machine humaine, la possibilité de coupler un très grand nombre de cerveaux, et l'aide éventuelle de l'informatique pour classer les données, on montre ainsi que la complexité du « cerveau analysant » n'est absolument pas limitée par celle du « cerveau analysé », ce qui élimine l'objection de départ de F. Jacob.

1. Les suites aléatoires et les démonstrations mathématiques, par Gregory Chaitin, in *Les progrès des mathématiques*, op. cit.

3. *La machine à penser de Turing*

JPC : Venons-en à la machine de Turing.

AC : A toi de me rappeler ce que c'est.

JPC : Turing fut un mathématicien exceptionnel. Ses travaux inspirent encore beaucoup de biologistes. Il fut ainsi un des rares créateurs en mathématiques à proposer des théories dont l'application en biologie a été décisive. C'est le cas, par exemple, pour sa théorie de la morphogenèse par ruptures de symétrie. Il a réussi à montrer comment, sur la base d'un système de réactions chimiques couplées, peut se créer spontanément une forme à partir d'un système isotrope. Il s'est d'ailleurs posé le problème de manière très concrète et amusante, en s'efforçant d'expliquer comment peut se former, à partir d'un œuf sphérique, une hydre la bouche entourée de six tentacules ! Des problèmes biologiques très concrets et précis peuvent, on le voit, inspirer des théories mathématiques originales. Mais il a aussi été l'un des premiers à formuler la théorie des machines informatiques, des ordinateurs tels qu'on les utilise aujourd'hui. Cette théorie fait toujours l'objet d'un débat très vif entre psychologues et neurobiologistes, et pose le problème de savoir si on pourra jamais construire une machine de Turing aux performances identiques à celles du cerveau humain, et si le cerveau humain est lui-même une machine de Turing. L'article de ce dernier commence d'ailleurs par cette phrase : « Je propose de réfléchir à la question : les machines peuvent-elles penser ? » C'est précisément la question que nous nous posons l'un et l'autre.

Mais d'abord, qu'est-ce qu'une machine de Turing ? Telle qu'il la décrit dans son article de 1936, celle-ci lit et

écrit des symboles discrets, des carrés, sur une bande qui sert d'entrée à la machine. Elle stocke des symboles et sert de mémoire. Mais elle sert également de sortie. La machine effectue trois opérations : elle lit les symboles, elle substitue et elle ajoute des symboles les uns aux autres. C'est une bande qui, en théorie, n'a pas de limites et qui définit en quelque sorte le programme. D'emblée, Turing distingue le *programme*, ou « *software* »...

AC : Est-ce qu'elle relit ce qu'elle a fait ?

JPC : Oui, elle en est capable.

AC : La bande passe-t-elle une seule fois, ou revient-elle ?

JPC : Elle peut repasser indéfiniment. Elle contient le programme, ou « *software* », tandis que le reste de la machine, « en dur », constitue le « *hardware* ». Nous voilà en face d'un ordinateur tel qu'on en construit aujourd'hui.

AC : Sans donner le mécanisme que la machine utilise.

JPC : C'est le problème. Cette machine est un calculateur *numérique* qui utilise des quantités sous forme discontinue. Elle diffère en cela des calculateurs *analogiques* qui mesurent des quantités physiques. Un calculateur numérique – et c'est un point très important de la théorie de Turing – est capable d'imiter n'importe quelle autre machine opérant sur des quantités discontinues. C'est donc une machine universelle, puisque n'importe quel processus est susceptible d'être représenté sous la forme d'une série d'instructions permettant la manipulation d'éléments discontinus. N'importe quel processus de ce genre peut donc, en principe, être reproduit par une machine de Turing. Et même un calculateur analogique peut être simulé par un calculateur numérique.

La question se pose maintenant de la validité des thèses, dites de Church et Turing, selon lesquelles ce qui peut être calculé par un être humain peut également l'être par une machine, ce qui peut être calculé par une machine peut

également l'être dans un programme récursif général ou partiel, et enfin ce qui peut être calculé par un être humain peut également l'être par ce programme. Cela revient à dire qu'on doit pouvoir identifier le cerveau et son fonctionnement à une machine de Turing. La doctrine du fonctionnalisme, très prisée des psychologues cognitifs comme Johnson-Laird pose que la psychologie se réduit à l'étude des programmes et qu'elle est donc indépendante de la neurophysiologie, puisque celle-ci étudie, quant à elle, la machine et son code. Tout ce qui concerne le psychisme fait donc partie du *software*, tandis que le cerveau, avec ses neurones et ses synapses, constitue le *hardware*. Il présente donc peu d'intérêt pour les fonctionnalistes, qui vont même jusqu'à conclure que la nature physique du cerveau n'exerce « aucune contrainte sur l'organisation de la pensée »¹. Selon cette doctrine à la mode dans le domaine des sciences cognitives, peu importe que le cerveau soit formé de protéines ou de silicone, peu importe le nombre et la nature de ses neurones. Seuls comptent les algorithmes auxquels s'identifient les fonctions cérébrales. S'intéresser au substrat neurobiologique, c'est perdre son temps!

4. La théorie de la matrice S en physique, homologue du fonctionnalisme en psychologie?

AC : On peut faire le parallèle entre les deux attitudes que tu viens d'opposer et leurs homologues en théorie des champs. Cette théorie s'efforce de comprendre le mécanisme d'interaction des particules élémentaires. Deux tendances s'opposent.

1. Johnson-Laird, *Mental Models*, loc. cit.

JPC : Peux-tu d'abord expliquer ce qu'est la théorie des champs ?

AC : Lorsqu'on fait de la mécanique quantique et qu'on essaie de la rendre relativiste, on s'aperçoit que des particules se créent et s'annihilent automatiquement. Le nombre de particules, contrairement à ce qui se passe en chimie, n'est pas constant. Il est donc nécessaire, pour aborder même des phénomènes très simples, d'étudier non pas des particules isolées, mais des champs, qui dépendent d'une infinité de variables. Cette théorie très complexe a connu un immense succès. Mais surtout l'analogie est criante entre la doctrine du fonctionnalisme et la théorie de la matrice S de Heisenberg. Peu importe, selon cette doctrine, ce qui se passe au moment des chocs de particules, seul compte la matrice qui fait passer de l'état initial du système, constitué, par exemple, d'une quinzaine de particules libres dont on connaît les impulsions et les masses, à l'état final, lui aussi présenté sous forme de particules libres. Cette matrice associe un nombre complexe à tout couple (i, f) , formé d'un état initial et d'un état final. La probabilité associée à ce couple est le carré du module de ce nombre complexe. La théorie propose d'analyser les propriétés de cette matrice sans que soit connu de manière précise le mécanisme qui régit les interactions à l'œuvre au moment des collisions. Comprendre la matrice S ne veut donc pas dire qu'on a compris ce qui se passe, mais qu'on dispose d'un modèle donnant des résultats adéquats à la réalité expérimentale.

JPC : C'est ce qu'on appelle une phénoménologie.

AC : Oui.

JPC : Les interactions ont lieu dans une « boîte noire », dont on ne se préoccupe pas. Les fonctionnalistes s'intéressent au cerveau de la même manière !

AC : Exactement. Cette démarche introduit un certain nombre de simplifications et de complications. On réussit ainsi à formuler les problèmes de manière plus simple parce

que les détails du mécanisme sont laissés de côté. Mais le nombre de solutions possibles au problème posé est tellement grand qu'on ne s'y retrouve pas. Comme l'a bien montré le développement de la physique, cette théorie, prise isolément, ne suffit pas. Mais elle a son utilité si, tout en s'efforçant de comprendre les phénomènes, en théorie fondamentale, on conserve comme but le calcul de cette matrice S . On ne doit donc pas éliminer totalement ce point de vue. La physique a montré au contraire que, de temps en temps, il était riche d'enseignements. Aujourd'hui, la théorie des cordes est en vogue. Elle est précisément issue de la matrice S . Veneziano a trouvé une matrice S qui vérifie certaines propriétés importantes et qui permet de deviner ce qui doit se passer dans le mécanisme d'interaction. Cette théorie des cordes est du reste très étrange, car elle pourrait très bien disparaître du jour au lendemain et n'avoir aucune application. En fait, elle n'a pratiquement aucun contact avec l'expérience.

JPC : Pour le neurobiologiste que je suis, les thèses fonctionnalistes sont utiles parce qu'elles permettent de mieux définir une fonction. Dans le meilleur des cas, elles la présentent sous une forme quantifiée. C'est de la physiologie, en quelque sorte. On mesure la fonction « de l'extérieur », sans entrer dans le mécanisme « interne ».

AC : Exactement. Ce qu'on analyse, ce sont les productions de la « boîte noire », ses aptitudes.

JPC : Pour moi, c'est là une définition fonctionnelle du problème, qui est bonne. Le modèle neural qui sera proposé devra rendre compte de ces fonctions. Je suis donc tout à fait d'accord avec toi. Je ne nie pas l'intérêt d'une approche expérimentale qui quantifie les fonctions. Mais je suis opposé à une optique exclusive selon laquelle la description de la fonction constituerait une « explication » suffisante. On touche ici à un problème qui mérite attention.

Si les thèses fonctionnalistes étaient exactes, une fonc-

tion cérébrale s'identifierait à un algorithme mathématique, ou même à plusieurs. Mais peut-on identifier la réalité extérieure à des idéalités mathématiques? Ces idéalités décrivent-elles intégralement les phénomènes? Toi-même tu résistes à cette idée, puisque tu penses que les modèles mathématiques qu'utilise la physique ne donnent pas une représentation intégrale de la réalité physique, qu'ils ne l'épuisent pas. Le fonctionnalisme me paraît plutôt une méthode d'approche des fonctions cérébrales qu'une philosophie. Ses défenseurs se heurtent à un obstacle épistémologique considérable : un algorithme mathématique peut-il être identifié à une propriété physique du cerveau?

AC : Il est clair, en physique également, que se contenter de la matrice S , c'est faire un pas en arrière par rapport à la théorie des champs. Mais les fonctionnalistes ont assurément leur mot à dire quand il s'agit de préciser quels résultats expérimentaux sont reproductibles, quelles sont les quantités auxquelles on doit s'intéresser.

JPC : Exactement. Mais ils vont plus loin. Ils pensent par exemple que la description d'un raisonnement ou la construction d'une phrase par un algorithme de calcul et leur simulation par une machine de Turing suffit pour comprendre comment fonctionne la pensée.

AC : Nous avons déjà la réponse. Il suffit de se reporter aux trois niveaux dont nous avons déjà parlé. La faculté de reproduire une phrase appartient au premier niveau. Le mécanisme qui permet de la reproduire a été donné à l'avance. Mais savoir changer de stratégie en cas d'erreur est tout différent. Ce type de mécanisme relève manifestement du premier niveau. Si l'on croit avoir compris le cerveau parce qu'on a compris le premier niveau, il est évident qu'on fait une grave erreur. Même au premier niveau, la machine de Turing ne résout rien, car elle ne

prend pas en compte le problème de la complexité des algorithmes ¹.

JPC : Donc, nous enterrons le fonctionnalisme.

AC : Pas complètement. Il peut être utile pour préciser les quantités qu'il est bon d'étudier. Evaluer une théorie en termes fonctionnels peut être intéressant. Mais on ne peut pas se contenter d'une démarche fonctionnaliste.

JPC : Ce serait défendre un point de vue très conservateur. L'histoire a en effet montré qu'analyser des niveaux sous-jacents à celui dont on cherche à rendre compte, que pénétrer à l'intérieur de la « boîte noire », et la disséquer pour réduire, puis reconstruire, un processus physiologique, a fait systématiquement progresser la connaissance dans tous les domaines.

AC : C'est vrai aussi pour la théorie des champs.

5. *Le cerveau de l'homme est-il un ordinateur?*

JPC : Il est satisfaisant de constater notre totale convergence de vue à cet égard. Venons-en maintenant au dernier point : la différence entre le cerveau humain et les « machines à penser » actuelles. Les ordinateurs dont nous disposons sont très performants pour certaines catégories d'opérations. Par exemple, ils calculent extrêmement vite. Ils font des multiplications à dix chiffres en quelques secondes, ou même en quelques fractions de seconde. Mais ils sont évidemment très limités dans d'autres domaines. Par exemple, un ordinateur aura d'énormes difficultés à reconnaître un coque-licot dans une forêt ou un papillon dans la jungle, alors

1. L'efficacité des algorithmes, par H. R. Lewis et Ch. H. Papadimitriou, in *Les progrès des mathématiques*, op. cit.

que l'homme le fait instantanément! On souligne souvent également que les machines sont dépourvues d'« affectivité », de « corps »! Mais surtout, incapables d'anticipation, d'intentionnalité, elles ne peuvent construire leur programme sans « maître » extérieur. Leurs facultés d'auto-organisation sont très réduites, voire inexistantes. J'aimerais bien savoir ce que tu penses, toi qui pratiques les échecs avec, pour adversaire, une machine qui sait jouer aussi bien, sinon mieux que l'homme.

Les ordinateurs utilisés actuellement manquent, me semble-t-il, de deux propriétés que le cerveau humain possède. On peut tout d'abord remarquer que, dans le cerveau, le programme et la machine – pour employer le modèle de Turing – sont, dès les premiers stades du développement, très intriqués avec l'architecture connexionnelle. Il est difficile, sinon impossible, de définir un programme indépendant de la connectivité de la machine cérébrale. Les objets de sens se déposent progressivement dans la mémoire à long terme, au cours du développement. Le hardware se construit progressivement en fonction de la composition génétique de l'individu, mais aussi de l'interaction constante avec le monde extérieur. Mais surtout, propriété au centre de nos entretiens, le cerveau se comporte comme une machine évolutive. Il évolue, selon un modèle darwinien, simultanément à plusieurs niveaux et suivant plusieurs échelles de temps. Voilà ce qui, selon moi, différencie le cerveau des machines construites actuellement. Outre, bien sûr, l'intentionnalité, propriété liée à l'évolution et peu abordée parce que relevant du niveau d'organisation le plus élevé. Qu'est-ce qui, selon toi, différencie le cerveau humain des machines construites de nos jours? Et comment en concevoir une qui se rapprocherait du cerveau humain?

AC : Examinons tout d'abord le cas des machines qui jouent aux échecs. L'intentionnalité est alors très simple : gagner la partie. C'est une chose extrêmement simple à

définir. Définir une fonction d'évaluation qui estime à quel point on est proche de l'intention poursuivie pendant le jeu est relativement aisé. On peut donc construire une machine qui utilise une fonction d'évaluation déterminée par cette intentionnalité bien définie. Dans le cas du cerveau au contraire, l'intentionnalité change selon les problèmes qui se présentent. Le cerveau doit ainsi créer lui-même la fonction d'évaluation adéquate à une intentionnalité donnée. Plus précisément, il doit pouvoir apprécier si cette fonction d'évaluation est adaptée à l'intentionnalité donnée. Il doit donc, j'ignore comment, posséder une fonction d'évaluation de fonctions d'évaluation !

JPC : C'est ce qu'on peut appeler, avec Granger, la raison stratégique.

AC : Oui, mais je voulais établir une hiérarchie. D'une part, nous avons les fonctions d'évaluation. Une fonction d'évaluation peut être identifiée à un but. Se donner une intentionnalité revient un peu à se donner une fonction d'évaluation. Toutes les fonctions d'évaluation, certes, ne sont pas bonnes, parce que certaines correspondraient à des intentionnalités contradictoires, tandis que d'autres ne seraient adaptées à aucune intentionnalité. Mais on peut définir plus ou moins une intentionnalité comme une fonction d'évaluation cohérente. Dans une situation donnée, le cerveau doit pouvoir élaborer lui-même ce genre de fonction d'évaluation. Il doit donc être capable de créer, ou, tout du moins, de choisir parmi celles qui existent déjà. Et pour ce faire, il doit lui-même posséder une fonction d'évaluation établie une fois pour toutes qui lui permette de savoir si la fonction d'évaluation qu'il crée est adaptée au but qu'il poursuit.

JPC : Ce mécanisme suppose la mémoire.

AC : Effectivement, la mémoire, les expériences acquises. Le cerveau peut s'appuyer sur des analogies pour comparer la situation présente à celles qu'il a connues auparavant.

JPC : Il existe d'une part une mémoire génétique. L'organisme humain, tel qu'il est aujourd'hui, résulte de multiples générations d'organismes qui, auparavant, ont déjà vécu ce genre d'expérience. La réponse à un problème nouveau qui se présente est donc inscrite dans la mémoire des gènes. D'autre part, le cerveau est ouvert sur la réalité extérieure, et surtout, peut puiser dans la mémoire à long terme qui s'est déposée au cours de l'expérience postnatale.

AC : C'est au deuxième niveau que se pose le problème fondamental. Quel peut être le mécanisme qui permet au cerveau de choisir une fonction d'évaluation appropriée à son but ? Quels critères permettent le choix ? Tant que ne sera pas compris ce phénomène, on sera très loin du deuxième niveau, comme c'est le cas dans les machines qui existent actuellement.

JPC : C'est-à-dire qu'elles n'en sont même pas au troisième niveau.

AC : Elles n'en sont qu'au premier niveau. Elles permettent seulement de faire des additions ou des multiplications, même extrêmement compliquées, ou bien de jouer aux échecs. Mais la fonction d'évaluation, comme l'intentionnalité, est toujours donnée à l'avance. Aucune machine n'est aujourd'hui capable de construire elle-même la fonction d'évaluation adaptée à l'intentionnalité qu'on lui propose.

JPC : Les ordinateurs actuels ne sont même pas capables d'avoir des intentions.

AC : Non, puisqu'ils ne sont pas en interaction évolutive avec le monde physique. Malgré leur mémoire, ils n'ont pas de passé autre que celui que nous leur imposons. Ils sont non évolutifs. Il est certain que l'affectivité intervient dans ce phénomène. Quand on se donne un but, c'est pour se faire plaisir, à moins d'être masochiste !

JPC : Cette capacité de se faire plaisir est elle-même déterminée par notre passé évolutif. Si nous nous auto-

détruisions en nous faisant plaisir, il est certain que nous ne serions plus là !

AC : Bien sûr. Mais je pense que le mécanisme qui permet d'estimer si la fonction d'évaluation est appropriée au but suppose l'affectivité. Celle-ci est en effet nécessaire pour que l'on puisse apprécier ce qui s'est passé. L'adaptation de la fonction d'évaluation au but proposé ne peut se mesurer que par le plaisir ou le déplaisir qu'elle engendre. Imaginons par exemple un joueur d'échecs, qui, bien que capable de calculer comme un ordinateur, choisit une mauvaise fonction d'évaluation. Il est bien évident qu'il sera grandement frustré lorsqu'il constatera qu'il perd toutes les parties qu'il joue. Le choix d'une mauvaise fonction d'évaluation ne lui aura procuré que du déplaisir. Mais ce dernier n'apparaîtra qu'à la fin des parties, et pas plus tôt. Sa fonction d'évaluation inadaptée l'empêchera de comprendre, au cours du jeu, que sa position est mauvaise et qu'il est en train de perdre. Pourtant, au vu du résultat final, il comprendra l'inadéquation de sa fonction d'évaluation.

JPC : N'oublions pas que ce système d'évaluation interne (plaisir/déplaisir) est lui-même prédéterminé par le passé évolutif de l'espèce (*voir figure 33*). Ces affects sont déjà déterminés dans leur réactivité aux signaux du monde extérieur et du monde intérieur.

AC : De nos jours, les machines supposent toutes une intentionnalité prédéterminée. De ce fait, elles en restent au premier niveau.

JPC : Mais alors, comment construire des machines qui accèdent au deuxième niveau ?



Fig. 16. — *Cynopithecus niger*, au repos.

D'après nature, par M. Wolf.



Fig. 17. — Le même, caressé et exprimant sa satisfaction.

FIGURE 33

Expression des émotions chez le singe. *Les émotions ressenties et exprimées par l'homme ont un passé évolutif. Charles Darwin, dans l'Expression des émotions chez l'homme et chez les animaux, a analysé en détail les manifestations extérieures des émotions, en particulier sur le visage, chez l'homme, et montré que beaucoup d'entre elles se reconnaissent déjà chez l'animal, en particulier chez le singe. Cette figure est extraite de l'édition française de 1877 de ce texte.*

6. *Une machine qui souffre et s'auto-évalue*

AC : Je peux seulement formuler de manière précise le problème. Une machine de ce genre devrait être en interaction évolutive avec le monde extérieur. Elle devrait pouvoir créer automatiquement une fonction d'évaluation correspondant à un but donné de l'extérieur. Elle devrait donc évaluer elle-même la stratégie qu'elle utilise, et aboutir à une fonction d'évaluation apte, si la machine disposait d'une mémoire et d'une puissance de calcul suffisantes, à faire, par exemple, un bon joueur d'échecs.

JPC : Mais est-ce implémentable? Cette idée a déjà été formulée. Pourquoi ces machines n'existent-elles pas? L'obstacle est-il théorique ou pratique?

AC : Je l'ignore. Pour moi, le seul mécanisme qui, chez l'homme, permette d'accéder à ce deuxième niveau, c'est justement l'affectivité.

JPC : On peut imaginer une machine dont le plaisir est mesuré par une grandeur qui varie, avec un seuil, un niveau moyen, de sorte que la machine optimise...

AC : Reprenons le cas des échecs. Supposons que la machine ne dispose pas de la fonction d'évaluation qui permet de bien jouer aux échecs. Elle dispose de tous les mouvements possibles, des règles du jeu, et d'une grande puissance de calcul, mais elle n'a pas la volonté de gagner. Comment la lui inculquer? A chaque coup, un bon ordinateur évalue sa position, la marque sur une échelle et sélectionne parmi tous les coups possibles celui qui rend maximale la valeur de la fonction d'évaluation. La machine que nous essayons de concevoir, elle, ne dispose pas de cette fonction d'évaluation. Il faut trouver un système qui lui

permette de l'acquérir. Il faudrait faire en sorte qu'à la fin du jeu, lorsqu'elle perd ou qu'elle se trouve en mauvaise position, elle ressente une douleur...

JPC : Si elle ressent une douleur, tu as déjà résolu le problème.

AC : Pas encore. Elle réagit seulement au résultat final du jeu.

JPC : Mais si tu as déjà inscrit dans la machine le fait qu'elle souffre lorsqu'elle perd, tu as la réponse.

AC : Non, parce qu'elle souffre seulement à la fin de la partie.

JPC : Tu possèdes déjà un élément de la réponse.

AC : Un petit élément. Parce que l'évaluation a lieu seulement à la fin. C'est tout.

JPC : Il suffirait alors que la machine ait un peu de mémoire, qu'elle « comprenne » ce qu'il faut faire pour atteindre son but, et ensuite, qu'elle développe les stratégies requises. Il lui faudrait donc une expérience. Dire que la machine souffre en cas de défaite me paraît résoudre déjà une partie du problème.

AC : Si la machine souffrait chaque fois qu'elle joue mal, nous aurions gagné, nous aurions trouvé la fonction d'évaluation.

JPC : Je crois que ce ne serait possible qu'après un certain nombre d'expériences.

AC : Tu veux dire qu'elle construirait graduellement sa fonction d'évaluation en corrélant les coups, pris séparément, qu'elle a joués avec le résultat de chaque partie. Cela me paraît raisonnable.

JPC : Tu as dit quelque chose de très important : ta machine souffrirait en cas de défaite.

AC : C'est un début. La souffrance ressentie à la fin permet un début d'évaluation. La fonction d'évaluation ainsi créée attribuerait aux parties jouées un résultat positif en cas de victoire, négatif en cas de défaite. La machine pourrait

d'ailleurs également mémoriser les parties jouées par d'autres joueurs avec pour seule évaluation leur résultat final. Mais il faut bien comprendre qu'une partie d'échecs se joue localement. Si une partie comportait quarante coups pour chaque joueur, ce n'est certainement pas au dernier que la machine devrait commencer à réfléchir. C'est localement qu'elle devrait le faire. Lorsque nous nous donnons un but déterminé, nous n'attendons pas le terme ultime de nos actions pour évaluer la distance qui nous sépare encore de ce but. Nous restons constamment en éveil. Au fur et à mesure que nous avançons, nous optimisons localement notre conduite en fonction des événements passés. Notre machine serait stupide, si elle se contentait de dire « Je perds, je gagne, je perds, je gagne », sans en déduire de conséquences locales. La réflexion est donc, selon moi, un mécanisme qui permet de rassembler les résultats globaux des parties mémorisées pour créer une fonction d'évaluation locale. Au fur et à mesure que les coups se profilent, la mémoire se réfère à des parties perdues ou gagnées : la fonction d'évaluation est ainsi créée. Si nous parvenions à construire une machine disposant de ce mécanisme, nous pourrions modifier les règles du jeu et la faire jouer à nouveau afin d'observer si elle s'adapte. Ce serait un bon critère.

JPC : Cette machine posséderait un générateur d'hypothèses.

AC : Bien sûr. Mais ce genre de générateur existe déjà dans les ordinateurs actuels.

JPC : Que leur manque-t-il alors ?

AC : La fonction d'évaluation !

JPC : Approfondis la question de l'implémentation d'une fonction d'évaluation. C'est très important.

AC : Je vais faire une proposition précise, malheureusement très peu économique. Elle suffira pourtant à montrer qu'il existe des solutions. Supposons que l'ordinateur ait en mémoire un millier de parties d'échecs et, pour chacune,

comme seule évaluation du jeu de chaque joueur, le résultat final de la partie. Cette fonction primitive d'évaluation consisterait à dire « X a perdu » ou « X a gagné » à la fin de la partie. Elle est stupide pour la simple raison qu'elle n'est pas locale. Je vais maintenant définir une fonction d'évaluation sur l'ensemble des fonctions d'évaluation *locales*. Elle consisterait au contraire à comparer le résultat de la partie avec le score que la fonction d'évaluation locale considérée donne, au cours de la partie, à chacun des joueurs. S'il existe une corrélation entre le résultat final de la partie et le résultat de la fonction d'évaluation locale, celle-ci est bonne, sinon elle doit être rejetée. Comme notre ordinateur a en mémoire un très grand nombre de parties, le caractère grossier de cette estimation disparaît et permet d'évaluer chaque fonction d'évaluation locale. On a ainsi défini une fonction d'évaluation universelle qui permet de localiser la réflexion.

JPC : Nous sommes tout près de la conscience.

AC : Non, pas de la conscience, parce que nous sommes au deuxième niveau. Nous approchons de la réflexion.

JPC : De la réflexion consciente.

AC : Nous ne serions pas loin d'un nouveau modèle d'ordinateurs capable d'adaptation, s'il n'y avait le problème de la complexité, qui empêche de réaliser des machines, quand la complexité de l'algorithme est à croissance exponentielle.

JPC : Et le troisième niveau ?

AC : Alors là...

JPC : La notion de fonction d'évaluation d'une fonction d'évaluation est intéressante.

AC : C'est absolument nécessaire.

JPC : On peut également concevoir la « conscience » comme une sorte de perception du perçu.

AC : Que veux-tu dire ? Je me plaçais, quant à moi, seulement au plan de la réflexion.

JPC : Oui, mais d'une réflexion sur la réflexion. N'est-ce pas déjà une prise de conscience ?

AC : Non. Pour moi, la réflexion est déjà une fonction d'évaluation sur les fonctions d'évaluation locales.

JPC : Et tu ne vois pas de degré supplémentaire ?

AC : Non. En présence d'un ensemble de buts possibles, il faut être capable de créer sa propre fonction d'évaluation. On a donc besoin d'une fonction d'évaluation des fonctions d'évaluation locales, qui procède par comparaison avec l'expérience et le résultat final. Cela illustre un principe très important, celui de localité...

JPC : Je suis entièrement d'accord. Le neurobiologiste, lui aussi, s'intéresse aux activités locales de neurones.

AC : Au deuxième niveau, la réflexion est locale. Et il est vrai que c'est la pensée consciente qui réfléchit. Mais au troisième niveau, le mécanisme n'est plus le même.

JPC : Que veux-tu dire ?

AC : Au deuxième niveau, on peut adapter une stratégie à un but fixé. Au troisième, celui de la créativité véritable, le but lui-même n'est pas connu. Le propre de la créativité réside dans l'absence de but préalable.

JPC : Je ne suis pas tout à fait de cet avis. Le créateur choisit seulement parmi divers buts possibles. C'est un niveau d'intention supérieur.

AC : Souvent, lorsqu'on cherche à atteindre un but, il arrive qu'on découvre autre chose. L'essentiel est alors de reconnaître la nouveauté et l'harmonie propres à ce qu'on rencontre. Il ne s'agit plus alors de réflexion, mais presque de la création d'un nouveau but.

JPC : Une création occasionnelle, et non intentionnelle !

AC : Certes. Ce que j'ai avancé précédemment ne s'applique pas au troisième niveau. Je supposais la fin clairement définie. La machine ressentait un certain plaisir à gagner ou un certain déplaisir à perdre. J'ai donc montré comment adapter la fonction de sélection à un but bien

défini. Mais il est vrai qu'au troisième niveau, bien qu'un but occasionnel puisse être donné à l'avance, l'effort pour y parvenir débouche soudain sur la reconnaissance d'une harmonie qui en elle-même le modifie.

JPC : L'aléatoire joue un rôle beaucoup plus important.

AC : Je n'en suis pas entièrement convaincu. Des bifurcations ont bien sûr lieu, mais le troisième niveau se caractérise avant tout par la reconnaissance d'une harmonie...

JPC : Oui, mais il faut bien que quelque chose existe déjà. Il y a donc, par incubation et cheminement latéral, genèse de...

AC : A mon avis, autre chose intervient, la reconnaissance d'une harmonie. Cela n'appartient pas au deuxième niveau.

JPC : C'est un niveau supérieur.

AC : Le niveau de la réflexion est dépassé. L'harmonie est perçue, mais par un mécanisme qui n'est plus de l'ordre de la réflexion.

JPC : Un mécanisme d'intégration en quelque sorte. C'est la pierre qui nous manquait pour parfaire l'édifice...

AC : Par exemple, ou bien un processus qui fait résonner un ensemble de systèmes de neurones.

JPC : Dans le plaisir esthétique, on peut parler de mise en harmonie, ou en résonance, d'une activité du cortex frontal avec le système limbique.

AC : Peut-être...

JPC : Le plaisir lui aussi est très important dans l'illumination.

AC : Oui. Alors que le mécanisme de la réflexion fait intervenir le plaisir ou le déplaisir seulement au stade final pour élaborer une fonction de sélection, cette fois tout se passe différemment.

JPC : On pourrait tout de même imaginer une machine de ce genre.

AC : Je l'ignore. Nous en revenons à notre problème :

existe-t-il une harmonie préétablie, à laquelle l'homme est sensible parce qu'il vit dans ce monde harmonieux, ou bien crée-t-il lui-même l'harmonie? Découvrons-nous la réalité harmonieuse... ou bien créons-nous l'harmonie de la réalité?

JPC : Nous revenons à notre problème de départ. Mais désormais tu présentes les choses sous la forme d'une alternative! Ou bien il existe dans le monde une harmonie préétablie, et nous vivons alors dans un monde platonicien. Ou bien nous essayons seulement de favoriser la résonance harmonieuse du monde extérieur avec le monde intérieur que nous nous efforçons d'édifier.

VII

Questions d'éthique

1. A la recherche des bases naturelles de l'éthique

« En tant qu'une chose a de la conformité avec notre nature, elle nous est nécessairement bonne. »

Spinoza, *Ethique*, 31

JEAN-PIERRE CHANGEUX : Le développement des connaissances scientifiques, en biologie comme en mathématiques, pose de nouvelles questions d'éthique. La grande presse discute l'antagonisme science-morale. Plus rarement, on s'interroge sur les fondements des jugements moraux.

D'abord qu'est-ce que l'éthique? Depuis Kant, les philosophes ont tendance à séparer éthique et morale, pour donner à la première un statut privilégié. La morale porte sur les conduites individuelles. Elle rassemble les prescriptions qui règlent le comportement, à un moment donné de l'histoire d'une société. L'éthique, elle, a une visée plus

générale. Elle est considérée comme une discipline dont l'objet est d'élaborer les fondements des règles de conduite, de construire, en quelque sorte, une *théorie rationnelle du bien et du mal*¹.

Les problèmes d'éthique affectent directement le neurobiologiste. Dans son travail quotidien tout d'abord. Quand il s'agit d'explorer le cerveau de l'homme, tout n'est pas possible. Des limites sévères à l'expérimentation s'imposent.

Ces questions sont débattues au sein de comités d'éthique qui regroupent des personnalités scientifiques, ainsi que les représentants des diverses familles spirituelles, en particulier des grandes religions révélées. Les recommandations de cette « magistrature morale » donnent lieu, parfois, à des textes de loi, en général bien acceptés. Des concepts aussi difficiles à définir sur le plan juridique que ceux d'être vivant (le spermatozoïde humain est-il un être vivant?), de mort cérébrale (un électro-encéphalogramme plat indique-t-il la mort?), de personne humaine (*l'Homo erectus*, s'il vivait encore, aurait-il les mêmes droits que *l'Homo sapiens*?)... donnent lieu à débat. Un dialogue nouveau s'instaure entre sciences humaines et sciences biologiques.

Toutes ces questions conduisent inévitablement à s'interroger sur les fondements des prescriptions morales. S'agit-il d'une forme de consensus dogmatique fondé sur quelques principes métaphysiques communs aux religions, d'une sorte de « pacte » entre autorités religieuses? Ou, au contraire, est-ce l'expression du bon sens populaire, d'un désir collectif, conforme aux vœux d'une majorité, en dehors de toute référence explicite à la métaphysique? Peut-on imaginer que la recherche d'objectivité qui accompagne ces débats fasse accéder l'éthique au rang de science?

1. B. Edelman & M.A. Hermitte, *L'homme, la nature et le droit* (C. Bourgois éd., 1988).

Formé, comme je l'ai été, à l'école d'André Lwoff, Jacques Monod, François Jacob... je ne peux rester indifférent à cette réflexion. Même si certains rejettent l'idée que l'éthique repose *exclusivement* sur la connaissance objective, il me paraît aujourd'hui indispensable de s'y référer pour édifier une éthique. Il faut donc prendre comme point de départ pour toute réflexion de cette nature les données de l'anthropologie, de l'histoire des religions, du droit, de la psychologie cognitive, et donc des neurosciences. Il devient légitime, à la fois, d'agir en scientifique, de construire des modèles validables et éventuellement révisables et de se référer, sans cesse, aux résultats de la science. Cette base de discussion est plus sûre qu'un quelconque postulat métaphysique, ou qu'une croyance, en perpétuel conflit avec le sens commun ou avec les données de la physique la plus élémentaire.

J'irai même jusqu'à partager la réflexion de Jacques Monod : la quête incessante de la vérité, qui est le premier mobile de la science, constitue *de facto* une éthique. Et peut-être la plus respectée au fil de l'histoire, même si, par-ci par-là, les conduites de quelques scientifiques paraissent faire exception. Mais bien des philosophes ont déjà souligné la difficulté de la tâche. Ce libre examen demande un effort de recherche approfondie. Une fois de plus, une authentique ascèse ! Avons-nous le courage et la force nécessaires ? Les situations sont souvent contradictoires, et de nombreux motifs entrent en jeu : c'est une tâche très lourde que de rechercher les fondements objectifs de la morale et d'élaborer des règles morales sur la base d'une réflexion sur les données de la science contemporaine. Il est alors plus facile de se référer à des *a priori* transcendants qu'aux données de la science, elles-mêmes parfois fugaces. Elaborer des règles précises de conduite entraîne dans un Univers de connaissances et de réflexions qu'il devient chaque jour plus difficile de maîtriser. L'application des mathématiques que constitue

l'informatique – avec ses bases gigantesques de mémoire – nous aidera-t-elle dans cette démarche? Devrons-nous tôt ou tard demander aux ordinateurs de décider des jugements éthiques ¹? Voilà une question qui intéresse aussi le mathématicien.

Quoi qu'il en soit, il en est de l'éthique comme des mathématiques. Toi et moi sommes des représentants de l'espèce animale *Homo sapiens sapiens*. Nous possédons un cerveau qui détermine l'acceptation ou le refus des règles morales. Mais il les construit également dans un environnement social défini, à un moment précis de l'histoire culturelle de l'humanité. Tout scientifique qui refuse de succomber au clivage mental confortable du croyant, qui souhaite rester cohérent avec lui-même et qui s'efforce de rejeter toute référence à la métaphysique, devra tenter, dans sa réflexion, de rechercher les bases *naturelles* de l'éthique. Ce n'est, somme toute, que réactualiser la démarche des Lumières ² et de la Révolution française, avec le bénéfice considérable que peuvent nous procurer les résultats récents des neurosciences, des sciences cognitives et de l'anthropologie sociale.

Gunther Stent, biologiste moléculaire fort respecté, organisait dès 1978 une Conférence Dahlem intitulée *Morality as a biological phenomenon*, dans le but de s'attaquer aux présupposés idéologiques de la sociobiologie alors naissante, mais sans pour autant abandonner la recherche des bases biologiques de la morale. Bien au contraire. Ce débat de fond remonte en fait, comme souvent, à l'Antiquité grecque où déjà s'opposaient deux thèses proches de celles qui portent sur le fondement des mathématiques. La position idéaliste, illustrée par Platon, est simple : la conduite

1. D.C. Dennett, *The Intentional Stance* (MIT Press : Cambridge Mass. Ed., 1987).

2. Voir D. Diderot, 1775, *Entretien d'un philosophe avec la Maréchale de****.

morale doit être en harmonie avec des principes qui appartiennent au monde des Idées. Or celui-ci, nous en avons longuement débattu, contient également les lois mathématiques de l'Univers et de la connaissance. A l'opposé, Démocrite, Epicure, puis Lucrèce s'intéressent à l'homme comme espèce animale : pour eux, la sagesse consiste à se débarrasser de tout préjugé métaphysique responsable du malheur des hommes (*voir figure 5*). Deux points de vue opposés donc, difficilement conciliables.

Il va de soi, tu t'en doutes, que, dans ce dialogue à distance, j'adopterai une position naturaliste, dénuée de tout recours à une quelconque métaphysique. Je suis loin d'avoir construit un système cohérent. Je voudrais te faire part seulement de quelques réflexions préalables, en forme d'esquisses, et je prends le risque de te les livrer à peine ébauchées. J'espère que tu ne m'en tiendras pas rigueur.

Premier problème sur lequel il nous faudrait réfléchir ensemble : l'existence ou non d'une universalité des règles morales. Si on adopte le point de vue platonicien, il doit exister une universalité éthique, tout comme il existe une universalité mathématique. Tu devrais donc être platonicien en morale, comme tu l'es en mathématiques.

Or, comme le montre l'anthropologie, la diversité des cultures est considérable. Il existe des variations importantes dans le mode de pensée, dans l'organisation sociale et, de ce fait, dans le jugement éthique d'une culture à l'autre (*voir figure 34*). Il existe un « autrui culturel », qui reste incompréhensible, ou du moins relativement difficile, à comprendre par celui qui n'en fait pas partie. Les conflits récents entre sunnites et chiïtes (Iran-Irak), protestants et catholiques (Irlande), juifs et musulmans (Israël), hindouistes et bouddhistes (Sri-Lanka)... témoignent de cette impénétrabilité culturelle réciproque qu'entretiennent des religions qui ne méritent plus guère leur nom, puisqu'elles



FIGURE 34

La tour de Babel. Ce tableau peint au *xvii^e* siècle par le Lorrain François de Nome (appelé également « Monsu Desiderio ») reprend l'épisode célèbre de la Tour de Babel, qui, selon la mythologie biblique, aurait été élevée par les fils de Noé pour atteindre le Ciel. La diversité des langues de ceux qui participaient à sa construction les aurait empêchés d'atteindre ce but. Cette toile a été choisie pour illustrer le relativisme culturel. La diversité des cultures et des langues s'accompagne d'une diversité des morales souvent incompatibles entre elles et source inépuisable de conflits. (Repris de *Connaissance des Arts*, oct. 1957, entretien avec le docteur F. Sluys.)

divisent plus qu'elles ne « relient »¹. Ce *relativisme moral*², ou plutôt cette *relativité* des morales, concorde avec la diversité des langues, des représentations culturelles, des croyances, ou du droit. Dans la mesure où les prescriptions et les normes morales varient d'une collectivité à l'autre, il paraît difficile de définir un critère sûr, dont l'extériorité pourrait permettre d'évaluer la supériorité éthique de telle ou telle croyance, de tel ou tel comportement. Chaque culture défend, bec et ongle, que sa morale est, de toutes, la mieux fondée. Chacun est convaincu que sa morale est la plus « naturelle » ! C'est un concert d'aveuglement et d'intolérance réciproque... où chacun est persuadé d'être dans le vrai ! Les Japonais, qui sont devenus nos principaux partenaires scientifiques et économiques, vivent depuis des millénaires avec une éthique dont les fondements historiques sont totalement différents de ceux du judéo-christianisme qui prévaut en Occident. Pourquoi leur morale serait-elle inférieure ou supérieure à celle-là ? L'anthropologie sociale place le naturaliste devant une situation d'une extrême difficulté, puisqu'elle met plus aisément en relief la diversité que l'universalité des règles morales. Il paraît donc très difficile, après analyse des règles qui ont existé ou qui existent dans les multiples communautés culturelles, de dégager un « universel moral » précis.

2. Vie sociale et lobe frontal

JPC : Le principal « universel » qui émerge pourtant avec certitude, c'est l'existence même des morales et de la

1. Le mot « religion » dérive du latin « religere » qui signifie relier, unir.

2. Nicole Sindzingre, *Autres cultures, autres mondes ?* Autrement n° 93, *L'éthique corps et âme*, ch. 4, *Un nouveau prêt-à-penser ?*

réflexion éthique par-delà la diversité des cultures. C'est, comme nous l'apprend Kant, l'universalité de *l'exigence éthique* elle-même (voir figure 35). On peut, dans ces conditions, englober sous ce terme l'ensemble de règles d'interactions entre individus membres d'un groupe social. L'éthique procéderait ainsi de l'existence même du fait social. C'est déjà un premier point. Le naturaliste va lier l'existence des morales et de l'éthique à celle de la vie en société, même si celle-ci n'est pas propre à l'espèce humaine.

En effet, le lien social est, chez certaines espèces animales, les insectes par exemple, d'une solidité bien supérieure à ce qu'il est chez l'homme. Chacun sait, par exemple, que, chez l'abeille domestique, les ouvrières assurent obligatoirement l'alimentation de la reine reproductrice, mais sont elles-mêmes stériles. Certaines espèces de guêpes construisent des édifices collectifs d'une extrême complexité, avec une coordination et une efficacité bien supérieures à celles des hommes. Chez les mammifères et chez l'homme, aux conduites prosociales, se superposent des conduites qualifiées d'*antisociales*, comme le fait familial, le comportement territorial, ou l'agression intra-spécifique, qui entrent en compétition avec des intérêts plus généraux et qui concernent la survie de l'espèce. Toutefois, chez l'homme, les modalités d'expression du lien social présentent des propriétés singulières, du fait qu'il est, de toutes les espèces animales vivantes, celui qui possède les facultés cognitives les plus développées. Il y a, chez l'*Homo sapiens*, conjonction du social et du rationnel et recherche de leur *conciliation*.

Dans ces conditions, l'éthique s'intégrerait à la rationalité des obligations à l'égard d'autrui qui s'imposent à chacun au sein du groupe social. Elle définirait un ensemble de maximes réglant, conformément aux exigences de la raison, la coopération des membres de la communauté. Elle porterait donc, tout d'abord, sur les modalités de *communication* entre membres du groupe social, et non seulement



FIGURE 35

La Liberté et l'Égalité unies par la Nature. Gravure anonyme de la fin du XVIII^e siècle illustrant de manière allégorique les fondements naturels des deux concepts premiers de la Déclaration des Droits de l'Homme, l'Égalité, à gauche, avec pour attribut l'équerre, et la Liberté, à droite, identifiable par le bonnet phrygien. La Nature est représentée sous la forme d'une déesse assise aux multiples seins, coiffée de tours, ceinte de cornes d'abondance et vêtue d'une jupe décorée des signes du zodiaque. Celle-ci fait un geste d'union entre les deux personnages féminins allégoriques qui se serrent la main. (Musée Carnavalet, photo Bulloz.)

sur la reconnaissance des actes du locuteur, mais aussi sur celle de ses *intentions*, modèle de communication qualifié d'inférentiel par Grice (1957) ¹ ou Sperber et Wilson (1986) ². La recherche, parmi les multiples « avatars » des diverses morales, de leurs fondements universels nous conduit à considérer certaines facultés cognitives caractéristiques de l'espèce humaine. D'abord celles de se *représenter* l'autre, avec ses états émotionnels, ses intentions, ses projets à court et long termes, mais surtout de se le représenter, de manière réflexive, comme un autre soi-même, membre de la même espèce sociale. Ces facultés cognitives incluent la capacité d'élaborer une théorie des états mentaux d'autrui ³, voire une théorie des théories qu'il échafaude pour l'avenir. Elles comprennent également celles de se représenter l'organisation du groupe social, et la possibilité d'actualiser en son sein des « états mentaux » individuels.

Ces facultés mettent en jeu des « architectures neurales », qui incluent des niveaux d'organisation proches de ceux que tu as définis comme le deuxième et le troisième niveaux de la pratique des mathématiques. Or nous savons l'importance du lobe frontal dans la plupart de ces cas. Les recherches cliniques sur les lésions du lobe frontal mettent effectivement en évidence des perturbations de conduites sociales, voire du « sens moral ». Dans sa description du cas de Phinéas Gage, jeune ouvrier blessé au niveau du lobe frontal par une barre de mine, Harlow, dès 1869, signale que « l'équilibre, la balance pour ainsi dire, entre ses facultés intellectuelles et ses penchants instinctifs semblent détruits. Il est nerveux, irrespectueux, et jure souvent de la façon la plus grossière, ce qui n'était pas dans ses habitudes aupa-

1. H.P. Grice, *Logical Conversation-William James Lectures* (Harvard U.P. Cambridge Mass éd., 1986).

2. D. Sperber et D. Wilson, *Relevance* (Blackwell éd. Oxford, 1986).

3. Cette faculté semble déjà présente chez le chimpanzé : voir Premack, A. and Woodruff G. (1978), *Does the chimpanzee have a theory of mind?*, Behavioral and Brain Sciences n° 1, pp. 515-526.

ravant; il est à peine poli avec ses égaux; il supporte impatiemment la contrariété, et n'écoute pas les conseils des autres lorsqu'ils sont en opposition avec ses idées... ». Luria décrit également ¹ le cas d'un patient qui, devant le tableau de Klodt *Le dernier printemps*, représentant une jeune fille mourante assise dans un fauteuil, interprète la scène comme un mariage, à cause de la robe blanche de la jeune fille. Le malade frontal ne comprend plus les éléments émotionnels du tableau, il ne les situe pas correctement dans leur contexte social. Ce n'est donc pas un hasard si Luria qualifie le cortex frontal d'« organe de la civilisation ».

Le rôle du cortex frontal se distingue de celui que jouent d'autres régions du cerveau comme le cortex temporal. Geschwind ² a décrit un cas curieux d'épilepsie du lobe temporal, retrouvé également par Gazzaniga ³, qui entraîne une intensification des convictions religieuses (avec de curieux et inattendus passages d'un système de croyances à un autre), accompagné du désir d'écrire abondamment (hypergraphie) et du goût pour des pratiques sexuelles bizarres. Le cortex frontal est en interaction constante avec les autres aires du cortex cérébral. Il n'y a pas *un* « centre » cérébral de l'éthique. Mais des ensembles hiérarchisés et parallèles de neurones contribuent aux fonctions « cognitives » qui servent à élaborer l'éthique. Ces prédispositions neurales à l'éthique sont, dans leur ensemble, communes à l'espèce humaine. Elles font partie des traits qui distinguent l'homme d'autres espèces animales. Elles sont donc sujettes aux déterminismes génétiques qui délimitent la « nature humaine ». Ce qu'il y a d'universel dans l'éthique qui peut conduire à la définition des Droits de l'Espèce

1. Luria, p. 333, 1978.

2. N. Geschwind, *Behavioral change in temporal epilepsy*. (Archives of Neurology n° 34, 453, 1977.)

3. M. Gazzaniga, *Le cerveau social* (Laffont éd., Paris, 1985).

Humaine, il faut en rechercher l'origine dans l'expression du patrimoine génétique commun à l'humanité.

3. Le comportement prosocial de l'enfant et l'empreinte culturelle

JPC : Ces déterminants génétiques s'expriment, progressivement et successivement, au cours du développement embryonnaire et fœtal, lorsque se mettent en place les grandes lignes de l'architecture cérébrale et, en particulier, la primauté du cortex frontal. Dès la naissance, l'enfant interagit avec « l'autre ». Des conduites « prosociales » se développent, qui assurent son interaction harmonieuse avec les autres personnes de son environnement ¹. Dès trois mois, il effectue des échanges avec sa mère ou son père. Dès un an, l'enfant apprend à partager. Spontanément, il montre et offre ² des objets à d'autres personnes pour entrer en communication avec elles. A onze mois, il prend soin des autres. Il donne à boire et à manger à sa poupée, avec une nourriture imaginaire. A deux-trois ans, il établit une conversation. Il présente aussi, très tôt, des sentiments d'amitié et d'affection, avec pour signes caractéristiques, sourires et baisers. Il présente de l'intérêt pour l'autre, mais a parfois peur de l'étranger. Enfin, à partir de huit ans, se manifeste l'aptitude à « se mettre à la place d'autrui » ou sympathie. La capacité à participer aux émotions de l'autre apparaît très précocement. Elle est à la base de la représentation de l'autre comme soi-même, que j'ai déjà men-

1. H.L. Rheingold et D.F. Hay, *Prosocial behavior of the very young*, p. 105-124, in G.S. Stent, « Morality as a biological phenomenon », 1978.

2. H. Montagner, *L'attachement : les débuts de la tendresse*, Odile Jacob, Paris, 1988.

tionnée, de l'autre non seulement comme individu, mais comme individu sentant.

La sympathie, comme conduite dont le but est de soulager l'inconfort de l'autre, se manifeste entre dix-neuf et trente-six mois. Puis, apparaissent les notions d'obéissance et de responsabilité consciente. Entre neuf et douze mois, l'enfant suit les ordres de sa mère, et à dix-sept mois, il se donne des ordres à lui-même. Le jeune enfant devient progressivement capable de conduites d'aide et de coopération. Il participe avec un autre à une action commune, dans un but commun. Les attitudes prosociales connaissent donc un développement progressif qui suppose, très vraisemblablement, un nombre important de conduites innées.

Toutefois, dès la naissance, l'interaction de l'enfant avec son environnement physique et social laisse des traces qui vont signer l'individualité de l'adulte, tout autant sinon plus qu'une quelconque hétérogénéité génétique. Hubel et Wiesel¹ ont montré que le fait d'élever un chat ou un singe dans un environnement de barres verticales, alternativement claires et obscures (ou avec les paupières closes à un seul œil (C. Blakemore)), modifie nettement la spécificité fonctionnelle des neurones du cortex visuel. Et ces résultats sont probablement valables pour d'autres régions du cerveau, en particulier le cortex frontal². Chez l'embryon, l'activité neurale spontanée peut jouer un rôle considérable dans l'épigenèse du système nerveux. Le développement des performances cognitives et des états émotionnels de l'individu est, vraisemblablement, soumis à une importante épigénèse par sélection. Les croyances et des règles morales se fixent parallèlement à l'acquisition de la langue maternelle, selon des modalités, peut-être, analogues. Le cerveau

1. D. Hubel & T. Wiesel, *Functional Architecture of Macaque Monkey Visual Cortex*, Proc. Roy. Soc. London B., 198, 1-59, 1977.

2. P.S. Goldman-Rakic, *Development of cortical circuitry and cognitive functions*, Child Dev. n° 58, 642-691, 1987.

de l'enfant « s'imprègne » des règles morales, comme du langage propre à l'environnement familial et culturel dans lequel il est élevé. Celui-ci lui impose de manière *autoritariste*, voire totalitaire, une appartenance culturelle particulière qui le marquera pendant des décennies, et dont il se dégagera difficilement par la suite, voire pas du tout. Bien sûr, les bases neurocognitives de la fixation des croyances restent, pour la plupart, inconnues. Mais elles constituent un sujet de recherche passionnant.

Il apparaît, dès lors, essentiel de préciser les régularités dues à ces facultés qui sont imposées par le patrimoine génétique humain et qui, formant une sorte de « grammaire générative » de l'éthique, détermineraient les principales étapes du comportement prosocial. Il importe ensuite de les distinguer des règles, propres à une culture définie, qui contribuent à sa singularité. Mais cette distinction reste très difficile, par suite de l'intrication très profonde de ces deux composants au cours des étapes successives du développement. Quoi qu'il en soit, la frange de variabilité connexionnelle qui se trouve libérée du pouvoir des gènes par les modalités de croissance et de stabilisation des connexions synaptiques permet à une morale particulière de s'installer dans un milieu social défini, à un moment donné de son histoire.

4. *Les fonctions de la morale*

JPC : Les facteurs qui déterminent l'installation d'une morale dans une collectivité animale ont donné lieu à des théories contradictoires. Les thèses, très controversées, de certains sociobiologistes, comme E.O. Wilson, se fondent sur des recherches réalisées chez les insectes, guêpes ou abeilles,

dont les comportements sociaux sont soumis à un déterminisme génétique extrêmement strict. Le généticien Hamilton (1964) a montré, de manière théorique, qu'un gène qui détermine le comportement suicidaire d'un sujet peut se répandre dans la population et introduire dans celle-ci un comportement « altruiste », si le suicide sauve cinq frères ou sœurs ou dix petits-cousins. D'où l'idée que la fonction de la morale est non seulement d'assurer la survie de l'espèce sociale, mais également de servir à la propagation de gènes qui déterminent les conduites sociales, en particulier le comportement prosocial de l'enfant. Mais l'extension théorique de ces idées fait passer sans transition de l'insecte à l'homme. On lit par exemple, sous la plume de E.O. Wilson, que « l'encéphale n'a d'autre raison d'être que d'assurer la survie et la multiplication des gènes qui organisent sa formation » ou que « les règles de mariage seraient des « stratégies de transmission de gènes » ¹. On cite souvent, à l'appui de ces thèses, les divers interdits sur les mariages entre groupes religieux, ou encore le célibat imposé aux prêtres catholiques, au bénéfice des doctrines morales qui, en s'opposant aux méthodes contraceptives ou à l'interruption de grossesse, permettent d'avoir plus d'enfants,... et donc de propager les gènes de ceux qui adoptent ces doctrines! Il n'est pas exclu que des mécanismes de ce type soient intervenus au cours de l'évolution des insectes dont les conduites sont si rigidelement déterminées. Mais même dans ce cas, la démonstration n'est pas faite. Il en a, peut-être, été également de même lors de l'hominisation. Celle-ci pose en effet un problème redoutable au généticien des populations ². L'accroissement spectaculaire de complexité cérébrale qui s'observe de l'Australopithèque à l'*Homo sapiens* s'est réalisé en quelques millions d'années, et même

1. E.O. Wilson, *L'humaine nature*, p. 29, 1978.

2. Voir. J.-P. Changeux, *L'Homme neuronal*, Fayard, Paris, 1983.

moins, par des mécanismes génétiques encore totalement inconnus.

Personnellement, j'ai toujours adopté une attitude très critique vis-à-vis de thèses qui assignent une relation aussi simpliste entre gènes et conduites sociales, qui escamotent l'épigenèse et, surtout, qui oublient que, chez l'homme, un des traits majeurs de l'éthique est de concilier conduites sociales et raison, alors que celle-ci n'est pas développée chez les insectes ! D'autre part, il existe des exemples de prescriptions morales, des rituels, dont les conséquences sont opposées aux effets dont je viens de parler. Un des cas les plus spectaculaires est la pratique anthropophagique qui, en Nouvelle-Guinée, propage le kourou, une maladie provoquée par un virus lent qui entraîne des dégradations cérébrales très graves chez l'adulte. Plus généralement, il existe beaucoup d'exemples de cultures où les règles morales sont considérées comme génétiquement neutres. Ce qui va d'ailleurs dans le sens de la multiplicité des croyances et des règles morales. Il n'existe qu'une relation très indirecte, voire nulle, dans les sociétés actuelles, entre les règles morales, propres à une culture, et l'aptitude darwinienne à transmettre les gènes qui détermineraient ces règles. La fonction la plus évidente de la morale est « épigénétique ».

Même « neutre » sur le plan génétique et arbitraire dans ses prescriptions, on peut penser que la fonction de la morale, au niveau social, consiste à régler les interactions entre individus et, de ce fait, à participer à la survie de l'espèce. Mais cette régulation s'applique d'abord à celle de la collectivité culturelle, plus restreinte, à laquelle l'individu appartient. Mise en application de facultés éthiques plus universelles, la morale faciliterait la communication inférentielle entre individus d'un groupe culturel particulier. Elle permettrait une économie de temps dans l'actualisation des intentions en conduites. Supprimant une série de processus intermédiaires de raisonnement, devenue système de

droits et de devoirs, elle offrirait des « condensés de rationalité », qui réduisent le « fardeau de pensée », et offrent des réponses *toutes faites* à des conduites à venir.

5. Pour une morale naturelle, rationnelle et révisable

JPC : Une fois reconnues, d'une part, l'idée de bases neuro-cognitives à l'éthique et de leur universalité pour l'espèce, et, d'autre part, l'existence incontestable d'une relativité de la morale d'une culture à l'autre, se pose la question des principes qui, dans la pratique, serviront à l'élaboration des règles morales. Concrètement, sur quelles bases distinguer le bien du mal? Le philosophe américain Nagel ¹ a regroupé les diverses théories éthiques en deux ensembles distincts : les théories de type *déductif*, ou encore « autoritaristes » ², fondées sur des axiomes *a priori*, ou évidents en soi, du bien et du mal, et les théories de type *inductif*, dont l'impératif est de refuser tout *a priori* métaphysique ou idéologique. Selon Nagel, le type même des théories déductives est celle de Kant, dans son application à des règles d'action : « Agis uniquement d'après la maxime qui fait que tu peux vouloir en même temps qu'elle devienne une loi universelle. » C'est également le cas des thèses *utilitaristes*, comme celle de Bentham et de John Stuart Mill ³, qui prennent pour fondement de la morale, l'utilité, c'est-à-dire « le principe du plus grand bonheur qui soutient que les

1. T. Nagel, *Ethics as an autonomous theoretical subject*; p. 221-232, in G. S. Stent, *op. cit.*, 1978).

2. P. Jacob, *Qu'est-ce-que l'autoritarisme épistémologique?* L'âge de la science n° 2, 25-58, Odile Jacob, Paris, 1989.

3. J.S. Mill, *Utilitarianism*, II-4 1863.

actions sont bonnes dans la mesure où elles tendent à augmenter le bonheur (le plaisir et l'absence de douleur), mauvaises en tant qu'elles tendent à en produire le contraire ». Les idéologies utilitaristes, tout autant qu'universalistes, même si elles apparaissent, à première vue, justifiées, sont souvent, dans la pratique, en contradiction avec leurs principes. L'universalisme moral se heurte à la diversité des cultures; l'utilitarisme, au bonheur de l'individu contre celui de la collectivité. Les théories morales déductives se prêtent au fanatisme, au dogmatisme le plus absolu, à un autoritarisme sans limites. Elles conduisent l'individu à abdiquer devant des postulats théoriques qui prétendent défendre le « bonheur » de l'humanité!

Pour Spinoza : « Rien ne nous est connu comme certainement bon ou mauvais que ce qui nous conduit à *comprendre véritablement les choses*, ou ce qui peut nous en éloigner » (*Ethique*, 27). L'intérêt se déplace, à l'opposé des théories déductives, vers les théories inductives. Selon elles, les principes éthiques sont adoptés et révisés sur la base de leur plausibilité, et de leur capacité à expliquer des jugements plus particuliers. Elles prennent donc en compte l'évolution culturelle de la société, de la connaissance scientifique, des techniques et des cultures. J'adopterai, bien entendu, le point de vue inductif, qui me paraît le plus acceptable pour le scientifique, du fait de la possibilité qu'il reconnait d'une révision des normes morales, en fonction à la fois de l'apparition de nouveaux problèmes pratiques et du progrès des connaissances. Ce point de vue se rapproche de celui de la théorie de la justice de Rawls, qui commence à être connue en France. Très schématiquement, Rawls défend la méthode dite de *l'équilibre réflexif*. Les jugements se développent et sont soumis à des épreuves *a posteriori*, avec le souci de maintenir un maximum de cohérence interne et d'objectivité. Chaque jugement crée une pression de critiques et de justifications pour des change-

ments de principes. Si le système social est redistributif, s'il rectifie les infortunes résultant des contingences sociales ou naturelles, il en résulte une éthique fondée sur la critique des normes morales et leur révision incessante pour libérer de nouvelles formes de conduites. Personnellement, cette philosophie me séduit parce qu'on peut lui découvrir des bases « neurales » et parce que, se rapprochant de la démarche de la science, elle protège d'une forme de totalitarisme, conséquence ultime des théories éthiques déductives. C'est une philosophie sans prétention, une « éthique des petits pas », qui résout les problèmes tels qu'ils se présentent, progressivement, et qui ne se fonde pas sur des postulats *a priori*, totalement inapplicables.

Dans ces conditions, il ne s'agit plus de soumettre la science aux impératifs des croyances, à l'autoritarisme des dogmes révélés ou d'une quelconque idéologie, mais de développer une *critique des croyances*, des *idéologies* et des *normes morales*, en fonction du développement de la science, pour en dériver de nouvelles règles de conduite plus objectivement justifiées¹. Je pense personnellement que le modèle inférentiel de communication, de reconnaissance des intentions, avec évaluation de leur cohérence rationnelle, et du développement d'un *équilibre réflexif* au sein du groupe social, permet d'élaborer une éthique dynamique, une « morale ouverte », sur des bases « neurocognitives » naturelles, sans aucun recours à des présupposés métaphysiques.

1. Le dernier article du second projet de Déclaration des droits de Sieyès (juillet 1789) énonce déjà clairement cette proposition : « Art. 42. Un peuple a toujours le droit de *revoir* et de *réformer* sa constitution. Il est même bon de déterminer des époques fixes où cette révision aura lieu, quelle qu'en soit la nécessité. »

6. « L'élargissement de la sympathie » et la fonction esthétique

JPC : La science a pour vocation première de pourchasser, en permanence, l'irrationnel pour atteindre la connaissance objective. Elle élabore des représentations en accord avec les faits d'observation. Même s'il existe de sévères limites à cette quête d'objectivité, celles-ci ont des conséquences moins graves que la subjectivité des croyances. En dépit du caractère invérifiable de leur contenu, de leur implausibilité tant physique qu'historique, les croyances se maintiennent, et même se répandent. Leur caractère « fondamentaliste » s'accroît. Paradoxe, dans un monde où la connaissance objective ne cesse de progresser ! L'anthropologie sociale et l'histoire des religions soulignent, en dépit du caractère dogmatique et « non révisable » des systèmes de croyance, leur dimension évolutive. Aux religions, aux idéologies succèdent d'autres religions, d'autres idéologies, souvent de manière abrupte et conflictuelle. Les prescriptions morales et le droit fondés sur ces systèmes de croyances sont sans doute, comme le suggérait déjà Epicure, des *contrats*, mais leurs conséquences sont redoutables. Les systèmes de croyances constituent les fondements essentiels des préjugés racistes ¹. Ils créent des antagonismes entre groupes ethniques, aussi importants, sinon plus, que les différences de couleurs de peau ou de forme des yeux. Ils sont exploités à des fins idéologiques par les pouvoirs politiques. Alors, pourquoi persistent-ils avec tant de vigueur ? Parmi les « représentations culturelles » publiques, on peut considérer

1. P.A. Taguieff, *La force du préjugé*, 1988.

que les croyances constituent une catégorie particulière de représentation mentale : un état précis d'activité de cellules nerveuses, que le sujet est susceptible d'utiliser dans les interactions qu'il établit avec ses congénères. Ce serait une sorte de « modèle », construit à l'intérieur du cerveau sur des bases physiques, matérielles, biologiques. Ces modèles, nous l'avons déjà longuement discuté ensemble, sont distincts des modèles scientifiques. Ils ne visent pas à la mise à l'épreuve, à la vérification, et, en général, contredisent le sens commun le plus élémentaire. Néanmoins ils sont susceptibles de se propager de cerveau à cerveau, et de les « infecter » avec le caractère épidémique¹ d'une attaque virale ! Ce caractère invasif, cette lutte des croyances les unes avec les autres, rappelle le « struggle for life » darwinien. Mais elle n'a pas de conséquences nécessaires au niveau génétique, je l'ai déjà dit. De plus, elle se situe à un niveau d'organisation et selon une échelle de temps distincte de l'évolution des organismes. A ce stade de la connaissance, les conditions de « sélection » ou de « stabilisation » des croyances dans un milieu culturel ne peuvent être formulées que sous forme d'hypothèses. Evoquons-en quelques-unes.

Parmi celles-ci, on mentionnera leur fonction de « substitut » de l'explication scientifique, soit à un moment de l'histoire de la connaissance où des données objectives ne sont pas encore disponibles, soit par suite de difficultés d'accès à cette connaissance dans un milieu social défini, à cause d'une éducation et/ou d'une information insuffisante. Les divers mythes sur l'origine du monde matériel, des espèces animales ou de l'homme témoignent de ce refus de l'inconnaissable, prélude à une première « mise en ordre » taxinomique de l'Univers, « forme timide et balbutiante de la science »². Ces substituts de causalité ont suffi à un

1. D. Sperber, *Anthropology. I. Psychology: Towards an epidemiology of Representations: Man.* (N.S.); 20, 73-89; 1984.

2. C. Lévi-Strauss, *La pensée sauvage*, p. 16, 21 (Plon éd., Paris, 1962).

moment donné de l'histoire, ou suffisent encore, en l'absence de données plus aisément assimilables, pour combler d'inquiétantes lacunes du savoir. Former l'enfant à la critique rationnelle, et l'aider très tôt à s'informer des données de la science, constituent des avantages sélectifs importants dans cette lutte incessante des croyances entre elles, et des croyances contre la science.

Mais la stabilisation des croyances et des idéologies peut également résulter de leur utilisation par les diverses formes de pouvoir à la tête des collectivités humaines qui s'intercalent entre l'individu et l'espèce. Il semble que, lors de la différenciation des premiers *Homo habilis* dans les plaines d'Afrique ¹, les populations humaines comptaient quelque 100 000 individus répartis sur des milliers de kilomètres carrés. Nous en sommes loin, et cet énorme accroissement de population depuis les origines de l'homme a entraîné un clivage de la société en groupes sociaux dotés de cultures distinctes. Cette *ségrégation des cultures* apparaît déjà chez l'*Homo erectus*, à la suite, semble-t-il, de la domestication du feu (il y a 400 000 ans). Ces groupes, culturellement distincts, se rassemblent autour d'institutions et de pouvoirs qui fondent leur identité, au moins pour partie, sur les systèmes de croyances. Cette forme d'institutionnalisation des croyances et des idéologies reste encore vive de nos jours. Dans ces conditions, la critique des croyances s'étend à celle, plus difficile encore, des institutions. Socrate la paya de sa vie. La question est plus directement politique : elle porte sur les relations de la science avec le pouvoir. Or celles-ci, comme les relations de la morale ou du droit avec la politique, sont, d'emblée, *contradictoires*. L'homicide est condamné entre individus *dans* l'Etat, mais glorifié *entre* Etats. Dans certains Etats, il est justifié pour le blasphème,

1. H. de Lumley, *Origine et évolution de l'homme*. (Musée de l'Homme, Paris, 1984.)

mais pas pour la polygamie... La morale des Etats s'oppose souvent à la morale du citoyen, comme à celle de l'espèce. Et les Droits de l'Homme constituent un premier rempart de l'espèce contre les Etats, les croyances et les idéologies. La défense d'une morale révisable et non contradictoire rejoint la lutte politique.

La critique des croyances et des idéologies reprise par la révision des normes morales sur la base des données de la science ne suffit certes pas pour construire une morale qui se fonderait sur des faits « neurocognitifs » avec la rigueur de la méthode scientifique. La défense de la « personne humaine », ou celle de « l'individu », paraît bien insuffisante devant l'ampleur du problème et se trouve souvent dirigée *contre* la science plutôt qu'elle ne s'appuie sur elle. Adam Smith, philosophe écossais des Lumières, connu surtout pour ses travaux d'économie, fut l'un des premiers à s'inspirer d'Epicure pour donner à la morale une explication purement naturelle, pour la séculariser. Dans sa *Théorie des sentiments moraux* de 1759, Smith considère comme élément principal la *sympathie*, qui « fait que nous avons conscience de l'effet qu'un acte aurait produit sur nous, ce qui fait naître en nous l'accord, ou le désaccord, avec les sentiments qui ont provoqué cet acte »¹. Pour Smith, il ne s'agit pas d'une « faculté », comme nous l'avons déjà mentionné, mais d'un produit de la vie sociale qui s'est lentement développé au sein de l'humanité. Charles Darwin, dans *La descendance de l'homme* renouvelle cette thèse, et considère que la sympathie est « une partie essentielle de l'instinct social », distincte de l'amour, que l'homme partage avec d'autres espèces animales, et qu'elle est innée, produite par la sélection naturelle. Mais l'homme est « un être moral, capable de comparer ses actes ou motifs passés ou futurs et

1. Cité par P. Kropotkine, *L'éthique*, p. 239 (Stock éd., Paris).

de les approuver ou de les désapprouver ¹ ». « A mesure que l'homme avance en civilisation et que les petites tribus se réunissent en communautés plus nombreuses, la simple raison indique à chaque individu qu'il doit *étendre* ses instincts sociaux et sa sympathie à tous les membres de la même nation... à tous les hommes de toutes les nations et de toutes les races ²... « aux infirmes, aux idiots et aux autres membres inutiles de la société » ³. Comme le relève Patrick Tort ⁴, cette thèse de « l'élargissement de la sympathie » innocente Darwin des propos racistes et des partis pris inégalitaires qu'ont adoptés, après lui, certains de ses contemporains comme T.H. Huxley ⁵. Au tournant du siècle, en 1906, Kropotkine, dans un ouvrage passionnant, *L'Entraide, un facteur de l'évolution*, reprend, développe la position de Darwin, l'enrichit de multiples observations empruntées au monde animal ou aux sociétés humaines « primitives » et à leur histoire. Pour lui, « dans l'évolution du monde organisé, le soutien mutuel entre individus joue un rôle beaucoup plus important que leur lutte réciproque » ⁶. « Plus les individus s'unissent, plus ils se soutiennent individuellement, et plus grandes sont, pour l'espèce, les chances de survie et de progrès dans le développement intellectuel. » Kropotkine considère l'entraide comme « un instinct de solidarité et de sociabilité humaine » ⁷, sur lequel se fondent les « sentiments moraux supérieurs » que sont « justice et moralité », et qui « amène l'individu à considérer les droits de chaque autre individu comme *égaux* aux siens » ⁸.

Les développements plus récents des sciences cognitives

1. P. 119.

2. P. 132.

3. P. 134.

4. P. Tort, *La pensée hiérarchique et l'évolution*. (Aubier, Paris, 1983.)

5. T.H. Huxley, *Struggle for existence as its bearing upon man*. (1888).

6. P. Kropotkine, *L'entraide, un facteur de l'évolution*, p. 9 (Hachette, Paris, 1909.)

7. *Ibid.*, p. 12.

8. *Ibid.*, p. 13.

n'altèrent en rien les positions de Darwin ou de Kropotkine. J'ai mentionné que la sympathie, ainsi que les conduites d'aide et de coopération apparaissent dans le « comportement prosocial » de l'enfant à des moments définis du développement. Une logique du « bien et du mal » peut donc être envisagée, qui retienne comme bon ce qui « élargit la sympathie », facilite l'entraide, et comme mauvais ce qui la restreint, la rend difficile. Il ne s'agit pas de viser à une quelconque universalité, utilité ou réciprocité. Mais, face à des choix limités et restreints, qui conduisent à une incessante révision des normes, comme le souhaite Rawls, il devient légitime d'adopter, de *sélectionner* par la raison, le jugement qui favorise, au moins localement, l'entraide au détriment des luttes individuelles ou collectives, et qui fasse s'exprimer la disposition naturelle de la sympathie, elle-même facteur positif de l'évolution des sociétés humaines. La sympathie, qui fut sans doute un facteur important dans la formation des premiers groupes sociaux d'*Homo sapiens* sert désormais, par son « élargissement », à fonder une « morale de l'Espèce » qui brise les cloisonnements en groupes culturels distincts. Alors se réalise la *conciliation* souhaitée entre le social et le rationnel.

Cette sécularisation de la morale par l'évolutionnisme met en valeur un *espace de variabilité* ¹ qui la fait échapper à sa « prétendue origine sacrée » (Spencer), ainsi qu'à l'autoritarisme fixiste des dogmes et des idéologies. Nous avons déjà considéré ensemble l'intervention d'un « générateur de diversité » aux deuxième ou troisième niveaux de l'organisation de la matière cérébrale, lors de la création des objets mathématiques. Ce générateur peut bien entendu intervenir au niveau de la production des représentations mentales de prescriptions morales qui portent sur les conduites sociales ou individuelles. Ces variations darwiniennes des représen-

1. Cl. Lévi-Strauss, *Race et histoire*, UNESCO, 1952.

tations sociales seront susceptibles ensuite d'être ou non propagées de cerveau à cerveau, sélectionnées au niveau d'une collectivité, puis retenues par le législateur, par exemple, sur la base d'un élargissement de la sympathie ou de l'entraide. Cela me rappelle une phrase du *Manifeste différentialiste* d'Henri Lefebvre : Peut-on proposer le « droit à la différence » ? Je réponds bien entendu par l'affirmative en précisant : droit à la différence, c'est-à-dire accepter à la variabilité avec sa composante *aléatoire*. Car il ne peut y avoir évolution sans variabilité à tous les niveaux d'organisation que j'ai discutés, pas seulement bien sûr au niveau génétique ni à celui des connexions du cerveau en développement. Je dirais même qu'il ne peut y avoir « révolution » sans variation préexistante, que celle-ci porte sur la forme du corps ou sur les productions du cerveau, y compris les modèles de sociétés humaines qu'il conçoit. Bloquer tout processus de variabilité par une quelconque « dictature » serait, je pense, bloquer la fonction d'anticipation qui est propre au cerveau humain. Ce serait freiner sa capacité d'intégrer les données de son environnement culturel pour produire des modèles, des idées novatrices, qui contribuent à sa dynamique évolutive. Il est donc légitime d'accepter *la variation aléatoire* dans toute éthique naturelle qui se veut évolutive. N'est-ce pas là une des définitions les plus dynamiques qui soit de la *liberté* : le droit à l'imagination ?

La critique rationnelle des croyances et des idéologies, l'élargissement de la sympathie, et le droit à l'imagination, suffiront-ils pour édifier une éthique débarrassée de toute irrationalité ? Rien n'est moins sûr. Car aux croyances se trouvent souvent associés des états émotionnels qui « relient » entre eux les membres du groupe social. La rupture de ce « lien » crée un sentiment de détresse. Les bases neurales et la pharmacologie de cette émotion fondamentale de l'interaction sociale ont été étudiées sur des modèles ani-

maux¹. Le cri de séparation ou cri de détresse provoqué par l'isolement du petit ou de l'adulte est calmé par la morphine et accentué par les agents pharmacologiques qui bloquent sélectivement l'effet des opiacés². Mais les rites et les croyances peuvent apparaître, Lévi-Strauss nous l'enseigne, « comme autant d'expressions d'un acte de foi en une science à naître »³, comme une première « taxinomie » qui « possède une éminente valeur esthétique »⁴. Ce plaisir, commun à l'homme et aux animaux⁵, de classer les objets du monde qui les entoure se retrouve, à un niveau plus élaboré, dans la création mathématique, tu l'as souligné à plusieurs reprises, et dans la création scientifique en général. La connaissance scientifique entraîne un plaisir qui peut sérieusement entrer en compétition avec celui des croyances.

Mais l'art ne joue-t-il pas directement et pleinement ce rôle? S'il a servi comme véhicule des croyances et des idéologies, ne peut-il pas, en lieu et place de la religion, jouer comme une puissance d'unification, comme une « raison communicationnelle »⁶ universelle, qui transcende la diversité des cultures et consolide la sympathie par un authentique plaisir collectif qui intègre cette diversité au lieu d'en faire un facteur de division, comme les religions, par nature intolérantes... Cette utopie esthétique, Schiller, dans ses *Lettres sur l'éducation esthétique de l'homme* (1795), la présente comme « édifice d'une véritable liberté politique ». Car selon lui, « seule la communication esthétique unit la société ». L'art joue le rôle de catalyseur de « l'har-

1. P.D. McLean, *The Midline Frontolimbic Cortex and the Evolution of Crying and Laughter*. (The frontal lobe revisited, pp. 121-141; I. Perecman ed. IRBN. Press; 1987).

2. Ce n'est pas hasard si l'usage de cette drogue se développe dans des conditions où l'intégration sociale apparaît insuffisante.

3. C. Lévi-Strauss, *La pensée sauvage*, p. 19.

4. *Idem*, p. 21.

5. N.K. Humphrey, *Natural Aesthetics*. (1980).

6. J. Habermas, *Le discours philosophique de la modernité* (Gallimard, Paris, 1988).

monie dans la société » par la « réconciliation des lois de la raison avec les intérêts des sens ». J'ai déjà longuement discuté les bases neurales et culturelles plausibles du plaisir esthétique et de l'art ¹. Je prends alors la liberté de rêver, avec Schiller, à un « Etat esthétique universel » qui affranchirait l'homme des armes et de l'irrationnel. Mais comme l'écrivait Spinoza en terminant l'*Ethique* : « Tout ce qui est beau est aussi difficile que rare. »

Voilà quelques réflexions, quelque peu hâtives et désordonnées, d'un neurobiologiste sur l'éthique. Qu'en pense le mathématicien ?

7. *Ethique et mathématiques*

ALAIN CONNES : Je commencerai par répondre à ton interrogation. Tu dis : « Si l'on adopte le point de vue platonicien, il doit exister une universalité éthique, tout comme il existe une universalité mathématique. Tu devrais donc être platonicien en morale, comme tu l'es en mathématiques. »

Ma réponse est négative. Ma « croyance » en l'existence d'une réalité mathématique brute, source inépuisable d'information, comme je l'ai expliqué au chapitre VI à propos du théorème de Gödel, est le fruit d'une longue expérience personnelle, et non de la lecture de Platon dont je ne partage pas nécessairement les idées.

Je ne crois pas en une éthique universelle, et il me semble que les mathématiques ne confèrent aucune compétence particulière pour parler d'éthique en termes généraux. L'idée d'une éthique mathématique n'a guère de sens. Dès

1. J.-P. Changeux, *Raison et plaisir*, loc. cit., p. 158.

lors, j'approuve entièrement ta critique des théories éthiques déductives fondées sur des *a priori* qui sont en fait idéologiques. Elle prend toute sa pertinence dans l'exemple suivant, bien connu des mathématiciens. La cryptographie à clé révélée¹ permet, à partir de certains résultats mathématiques très élaborés, comme ceux de la théorie des nombres, de fabriquer des codes inviolables qui peuvent être utilisés par les services secrets. Le mathématicien ne peut donc se retrancher dans sa tour d'ivoire, et affirmer que la pureté de ses recherches empêche qu'elles donnent jamais lieu à ce style d'applications.

Une théorie éthique déductive qui postulerait que les mathématiques pures ne doivent avoir aucune application au-delà du domaine civil conduirait à une impasse. Le mathématicien, s'il voulait lui rester fidèle, devrait renoncer à travailler, ou, tout au moins, à publier ses résultats. Il entrerait alors en contradiction avec l'éthique *de facto* que constitue la « quête incessante de la vérité » selon J. Monod.

Avec un peu de recul, on s'aperçoit en fait que ces cryptogrammes à clé révélée sont également très utiles à la société. Ils deviennent essentiels pour protéger l'individu contre les indiscretions de l'informatique, en particulier l'utilisation et la falsification d'informations privées, comme les renseignements médicaux confidentiels. On peut même envisager qu'à l'avenir, la signature de chacun, pour être inviolable, soit de nature mathématique.

Le mathématicien, comme tout autre scientifique, doit donc rester vigilant, au cas par cas, quant aux applications possibles de sa discipline, sans pour autant adopter une éthique déductive qui le neutraliserait.

Telles sont, résumées en peu de mots, les réflexions que je m'autorise en ce qui concerne l'éthique. Ni tour

1. Les mathématiques de la cryptographie révélée, par M. Hellman, in *Les progrès des mathématiques*, op. cit.

d'ivoire, ni éthique déductive, mais esprit de responsabilité scientifique. Ce qui m'importe surtout en fait, c'est de faire partager à d'autres l'essentiel de ce qui habite la recherche mathématique, le sens que l'on peut donner à « la quête du vrai », et la joie intérieure que l'on peut éprouver à s'y abandonner. Mes remarques, au fil de ces entretiens, ne cherchaient rien d'autre.

Table des matières

Avant-propos	7
I. Les mathématiques et le cerveau.....	13
Présentations	15
La hiérarchie des sciences mise en cause.....	19
Invention ou découverte?.....	25
Les mathématiques ont une histoire.....	35
Les mathématiques ne sont-elles qu'un langage?	39
II. Platon matérialiste?	43
L'ascèse intellectuelle du matérialiste	45
Psychanalyse des mathématiques.....	53
Les objets mathématiques sont-ils des représentations culturelles comme les autres?.....	55
Le darwinisme des objets mathématiques	58
Les croyances en mathématiques.....	62
III. La nature habillée sur mesure	65
Les mathématiques constructivistes	67
L'efficacité surprenante des mathématiques.....	74

Einstein et les mathématiques.....	80
L'utilité des modèles mathématiques en biologie	87
Auscultation de la mécanique quantique.....	95
IV. Le mathématicien neuronal.....	107
L'illumination	109
Le cerveau et ses multiples niveaux d'organisation	118
Le niveau cellulaire	127
Des circuits élémentaires aux objets mentaux.....	133
Neuropsychologie des mathématiques.....	143
La transition d'un niveau à l'autre par variation- sélection	146
Darwinisme mental et création mathématique	154
V. Darwin chez les mathématiciens	163
L'utilité du schéma darwinien.....	165
Coder des formes stables	170
L'organisation de la mémoire à long terme.....	184
Le raisonnement par analogie	188
Enchaînements de représentations et cadres de pensée.....	190
	194
VI. Les machines à penser	201
Des machines intelligentes?.....	203
Le théorème de Gödel	205
La machine à penser de Turing.....	214
La théorie de la matrice S en physique, homologue du fonctionnalisme en psychologie?.....	216
Le cerveau de l'homme est-il un ordinateur?.....	220
Une machine qui souffre et s'auto-évalue.....	226
VII. Questions d'éthique.....	233
A la recherche des bases naturelles de l'éthique	235
Vie sociale et lobe frontal	241
Le comportement prosocial de l'enfant et l'empreinte culturelle.....	246
Les fonctions de la morale.....	248

Pour une morale naturelle, rationnelle et révisable.....	251
« L'élargissement de la sympathie » et la fonction esthétique.....	254
Ethique et mathématiques	262

JEAN-PIERRE CHANGEUX ALAIN CONNES

Photo Louis Monnier



Quel est le lien entre le monde physique et le cerveau ? Les objets mathématiques existent-ils indépendamment du cerveau de l'homme, ou sont-ils seulement le produit de l'activité cérébrale ? L'éthique peut-elle être fondée sur des principes aussi universels que ceux des mathématiques ? Telles sont quelques-unes des questions essentielles instruites dans ce livre original et stimulant. Par la qualité de ses auteurs et la fécondité de leur réflexion, il constitue un événement exceptionnel.

Alain Connes, médaille Fields, est titulaire de la chaire d'Analyse et Géométrie au Collège de France.

Jean-Pierre Changeux, auteur de *L'homme neuronal*, est titulaire de la chaire de Communications cellulaires au Collège de France.

Tous deux sont membres de l'Académie des Sciences.

MATIÈRE À PENSÉE

EN ASSOCIATION AVEC LES ÉDITIONS DU SEUIL



9 782738 100733

ISBN 2-7381-0073-2

Imprimé en France

octobre 89

120 F